

ASTES

Advances in Science, Technology & Engineering Systems Journal

Special Issue

Multidisciplinary Sciences and Engineering

2019-20

www.astesj.com
ISSN: 2415-6698

EDITORIAL BOARD
(Special Issue)

Editor-in-Chief

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

Guest Editors

Prof. Wang Xiu Ying
Chongqing University, China

Prof. Yu Xiao Yan
Chongqing Normal University, China

Prof. María Jesús Espinosa Trujillo
Universidad Tecnológica Metropolitana, Mexico

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

Prof. Mohamed Abdelaziz Hassan Eleiwa
University of Hail, KSA

Prof. Nicolae Tudoroiu
John Abbott College, Canada

Editorial

The Special Issue on *Multidisciplinary Sciences and Engineering (2019–2020)* in the *Advances in Science, Technology and Engineering Systems Journal (ASTES Journal)* presents a broad and insightful collection of research that reflects the increasing reliance on multidisciplinary approaches to address complex scientific and engineering challenges. The period of 2019–2020 has been characterized by rapid technological progress alongside growing global demands that require integrated solutions spanning multiple domains. This special issue captures these developments by assembling contributions that bridge traditional disciplinary boundaries and promote the synthesis of knowledge across diverse fields.

A central theme throughout this issue is the integration of scientific inquiry with engineering innovation. The contributions demonstrate how principles from physics, chemistry, mathematics, and computational sciences can be effectively combined with engineering methodologies to develop robust and efficient solutions. This convergence has enabled the creation of more accurate models, advanced analytical techniques, and improved system designs capable of addressing multifaceted problems. By embracing multidisciplinary perspectives, researchers are better equipped to tackle challenges that extend beyond the scope of any single discipline.

The role of emerging technologies in facilitating multidisciplinary research is also prominently highlighted. Advances in data analytics, artificial intelligence, and high-performance computing have significantly enhanced the ability of researchers to collaborate and innovate across fields. These technologies support the integration of large and diverse datasets, enabling more comprehensive analysis and informed decision-making. The studies included in this issue illustrate how such tools are being applied in areas such as healthcare systems, environmental monitoring, smart infrastructure, and advanced manufacturing, demonstrating the far-reaching impact of multidisciplinary approaches.

Another important aspect addressed in this collection is the focus on sustainability and societal impact. The complexity of contemporary global challenges, including climate change, resource management, and public health, necessitates solutions that are both technically sound and socially responsible. The research presented emphasizes the development of systems that are efficient, resilient, and environmentally sustainable. By combining insights from multiple disciplines, these contributions offer innovative strategies for achieving long-term sustainability and improving quality of life.

Interdisciplinary collaboration and knowledge exchange are further emphasized as essential drivers of innovation. The works featured in this issue highlight the importance of collaborative research frameworks that encourage interaction among experts from different domains. Such collaboration fosters creativity, accelerates problem-solving, and leads to the development of integrated systems that are more adaptable and effective. This approach is particularly evident in the advancement of cyber-physical systems and intelligent environments, where seamless interaction between physical and digital components is critical.

The editorial team extends its sincere appreciation to all authors, reviewers, and contributors for their dedication and scholarly efforts in making this special issue possible. Their contributions have ensured the publication of high-quality research that advances the frontiers of multidisciplinary sciences and engineering. The collection presented in this issue reflects the ongoing evolution of research toward more integrated and collaborative paradigms, providing

valuable insights and a strong foundation for future exploration and innovation across diverse scientific and engineering domains.

Guest Editor

Prof. Yu Xiao Yan

ADVANCES IN SCIENCE, TECHNOLOGY AND ENGINEERING SYSTEMS JOURNAL

Special Issue **August 2019**

CONTENTS

Single-leg Standing Ability and Lower Limb Movement Analysis of Collegiate Footballers and Sedentary Students
by Jiun Sien Lau, Vina Phei Sean Tan, Mohd Akmal Vahidi, Siti Hanisah Linggi Mohd Azam, Mohd Hafezi Mat Zain, Azilawati Azizan, Mazra Othman and Rosniwati Ghafar

ANFIS-Based Climate Controller for Computerized Greenhouse System
by Hassan Oubehar, Abdelouahed Selmani, Abdelali Ed-Dahhak, Abdeslam Lachhab, Moulay El Hassane Archidi and Benachir Bouchikhi

A Proposal of TCP Fairness Control Method for Two-Host Concurrent Communications in Elastic WLAN System Using Raspberry Pi Access-Point
by Rahardhita Widyatra Sudibyo, Nobuo Funabiki, Minoru Kuribayashi, Kwenga Ismael Munene, Md. Manowarul Islam and Wen-Chung Kao

A Novel Hybrid Method for Segmentation and Analysis of Brain MRI for Tumor Diagnosis
by Kapil Kumar Gupta, Namrata Dhanda and Upendra Kumar

An ML-optimized dRRM Solution for IEEE 802.11 Enterprise Wlan Networks
by Mehdi Guessous and Lahbib Zenkouar

Distributed Linear Summing in Wireless Sensor Networks with Implemented Stopping Criteria
by Martin Kenyeres and Jozef Kenyeres

Centralized System of Universities Learning Materials
by Ruslan Vynokurov, Volodymyr Tigariev, Oleksii Lopakov, Kateryna Kirkopulo and Olena Pavlyshko

Ontological and Epistemological Issues of Studying the Development of Uzbekistan in the Context of Globalization from the Point of View of the Paradigm of Civilizations
by Djurayeva Nigora Avazovna

Dynamic Objects Parameter Estimation Program for ARM Processors Based Adaptive Controllers
by Vasiliy Olonichev, Boris Staroverov and Maxim Smirnov

Design and Validation of a Meter Band Rate in OpenFlow and OpenDaylight for Optimizing QoS
by Mohamed Saleh Al Breiki, Suiping Zhou and Yuan Roger Luo

Improvement of the Natural Self-Timed Circuit Tolerance to Short-Term Soft Errors
by Yuri Afanasyevich Stepchenkov, Anton Nikolaevich Kamenskih, Yuri Georgievich Diachenko, Yuri Vladimirovich Rogdestvenski and Denis Yuryevich Diachenko

Performance Evaluation and Examination of Data Transfer: A Review of Parallel Migration in Cloud Storage
by Mimouna Alkhonaini and Hoda El-Sayed

A Regenerative Energy Recovery System for Electric Vehicles Charging A Battery at A Low Speed
by Wonseok Yeo, Sungchul Jung, Seungtaik Kim, Keonho Park and Jongsun Ko

Multi Biometric Thermal Face Recognition Using FWT and LDA Feature Extraction Methods with RBM DBN and FFNN Classifier Algorithms
by Evan Hurwitz and Chigozie Orji

In Body Antenna for Monitoring and Controlling Pacemaker
by Sourav Sinha, Ta-Seen Reaz Niloy, Raja Rashidul Hasan and Md. Abdur Rahman

Routing Protocols for VANETs: A Taxonomy, Evaluation and Analysis
by Amina Bengag, Asmae Bengag and Mohamed Elboukhari

Attacks Classification and a Novel IDS for Detecting Jamming Attack in WBAN
by Asmae Bengag, Amina Bengag and Omar Moussaoui

Modelling the Methodology to Assess the Effectiveness of Distributed Information Systems
by Ivan Aleksandrovich Tarkhanov, Galina Pavlovna Akimova, Matvey Aleksandrovich Pashkin and Aleksander Vladimirovich Soloviev

A Harmonized European Drone Market? – New EU Rules on Unmanned Aircraft Systems
by Anna Konert and Tadeusz Dunin

Constrained Path Planning for Both Smooth Motion Profile and Stable Control
by Anh Son Tran and Ha Quang Thinh Ngo

A Word Spotting Method for Arabic Manuscripts Based on Speeded Up Robust Features Technique
by Noureddine El Makhfi

Implement Wireless and Distributed Vibrator for Enhancing Physical Activity of Visually Impaired Children
by Hung-Chi Chu, Fang-Lin Chao, Liza Lee and Pei-Yun Kao

New Color Image Encryption for Medical Images Based on Three Dimensional Generalized Chaotic Cat Map and Combined Cellular Automata
by Un Sook Choi, Sung Jin Cho and Sung Won Kang

Comparative Study of Semantic and Keyword Based Search Engines
by Bzar Khidir Hussan
Discriminant Analysis of Diminished Attentiveness State Due to Mental Fatigue by Using P300
by Kosuke Fujita, Kazuya Miyanaga, Fumiya Kinoshita and Hideaki Touyama

Model of Automatic Parking House

by Radek Guráš and Miroslav Mahdal

Internalising Negative Self-Image Externalities: The First Objective for City Marketing as a Municipal Management Tool
by Manuel Escourido Calvo, Antonio Javier Prado Domínguez and Valentín Alejandro Martínez Fernández

Learning Literary Style End-to-end with Artificial Neural Networks
by Ivan P. Yamshchikov and Alexey Tikhonov

Performance Portability and Unified Profiling for Finite Element Methods on Parallel Systems
by Vladyslav Kucher, Jens Hunloh and Sergei Gorlatch

Green Inventory Supply Chain Model with Inflation under Permissible Delay in Finite Planning Horizon
by Seema Saxena, Vikramjeet Singh, Rajesh Kumar Gupta, Nitin Kumar Mishra and Pushpinder Singh

The Potential of Ocean Current as Electrical Power Sources Alternatives in Karimunjawa Islands Indonesia
by Aris Ismanto, Dwi Haryo Ismunarti, Denny Nugroho Sugianto, Siti Maisyarah, Petrus Subardjo, Agus Anugroho Dwi Suryoputro and Hendry Siagian

Big Data Analytics Using Deep LSTM Networks: A Case Study for Weather Prediction
by Shweta Mittal and Om Prakash Sangwan

Spiral Curve for Revocable Touchless Fingerprint Template Securisation
by Tahirou Djara, Boris Sourou Zannou and Antoine Vianou

Design and Implementation of State-PID Feedback Controller for Poultry House System: Application for Winter Climate
by Ilyas Lahlouh, Ahmed Elakkary and Nacer Sefiani

The Acquiring Process of Entrepreneurship Competencies in Realist Airbrush Art
by Bayu Rahmat Setiadi, Suparmin, Slamet Priyanto, Henny Pratiwi and Sugiyono

A New Wire Optimization Approach for Power Reduction in Advanced Technology Nodes
by Jalal Benallal, Lekbir Cherif, Mohamed Chentouf, Mohammed Darmi, Rachid Elgouri and Nabil Hmina

Into the Online Space: Outcomes for a PASS Online Pilot Across Different First Year University Subjects
by Tracey Woolrych, Melissa Zaccagnini, Moira Stephens, Murray Stace, Rebekkha Middleton, Melissa Bergner, Sian O'sullivan and Reetu Verma

Coordination between Heterogeneous Models Using a Meta-model Composition Approach
by Naima Essadi and Adil Anwar

Non Parallelism and Cayley-Menger Determinant in Submerged Localization

by Anisur Rahman

SiC-FET Gas Sensor for Detecting Sub-ppm Gas Concentrations
by Yoshitaka Sasago, Hitoshi Nakamura, Takahiro Odaka, Atsushi Isobe, Shigenobu Komatsu, Yohei Nakamura, Taizo Yamawaki, Chiko Yorita, Nobuyuki Ushifusa, Kohei Yoshikawa, Kazuo Ono, Yumiko Anzai, Shuntaro Machida, Masaharu Kinoshita, Koji Fujisaki, Kenji Okishiro and Yuta Sugiyama

Transfer Learning and Fine Tuning in Breast Mammogram Abnormalities Classification on CBIS-DDSM Database
by Lenin G. Falconi, Maria Perez, Wilbert G. Aguilar and Aura Conci

Analysis and Improvement of an Innovative Solution Through Risk Reduction: Application to Home Care for the Elderly
by Linda Acosta-Salgado, Auguste Rakotondranaivo and Eric Bonjour

Testing Web Service Compositions: Approaches, Methodology and Automation
by Dessislava Petrova-Antonova, Denitsa Manova and Sylvia Ilieva

Chatting with Plants (Orchids) in Automated Smart Farming using IoT, Fuzzy Logic and Chatbot
by Samruan Wiangsamut, Phatthanaphong Chomphuwiset and Suchart Khummanee

A Study on Development of Information and Communication Ethics Sensitivity Measurement for Elementary School Students
by Woochun Jun

DFIG Defects Diagnosis Method for Wind Energy Conversion Chain
by Fatima El Hammouchi, Lamia El Menzhi and Abdallah Saad

An Algorithm to Improve Data Accuracy of PMs Concentration Measured with IoT Devices
by Mihaela Balanescu, George Suciu, Marius-Alexandru Dobrea, Cristina Balaceanu, Radu-Ioan Ciobanu, Ciprian Dobre, Andrei-Cristian Birdici, Andreea Badicu, Iulia Oprea and Adrian Pasat

Configurable Process Model: Discovery Approach from Event Logs
by Rabab Sikal, Hanae Sbai and Laila Kjiri

Hybrid Solar Thermal/Electricity Automated Oven
by Ajah Onu Victor and Emenike Ejiogu
Biometric System Vulnerabilities: A Typology of Metadata
by Abdou-Aziz Sobabe, Tahirou Djara and Antoine Vianou

An Algorithm for Automatic Measurement of KI-67 Proliferation Index in Digital Images of Breast Tissue
by Guillermo Kemper, David Atencia, Ivan Ortega, Roberto Kemper and Alejandro Yabar

System of Sequential D-Optimal Identification for Dynamic Objects in Real-Time Scale
by Fatuev Victor and Mishin Anton

Alternative Financing Model for Smart Cities Initiatives in Indonesia
by Robertus Nugroho Perwiro Atmojo, Tota Pirdo Kasih and Yakob Utama Chandra

Enhancing an SDN Architecture with DoS Attack Detection Mechanisms
by Ranyelson Neres Carvalho, Lucas R. Costa, Jacir Luiz Bordim and Eduardo Adilo Pelinson Alchieri

Development of Evaluation Metrics for Learners in Unplugged Activity
by Woochun Jun

SCMS: Tool for Assessing a Novel Taxonomy of Complexity Metrics for any Java Project at the Class and Method Levels based on Statement Level Metrics
by Issar Arab, Bouchaib Falah and Kenneth Magel

Stability of Ionized Air Using Density Functional Theory (DFT) for High Voltage Engineering
by Fri Murdiya and Neni Frimayanti

Framework for Automation of Cloud-Application Testing using Selenium (FACTS)
by Md Nurul Islam and Syed Mohammad Khurshaid Quadri

Attacks classification and security mechanisms in Wireless Sensor Networks
by Amine Kardi and Rachid Zagrouba

Current Trends and Challenges in Link Prediction Methods in Dynamic Social Networks: A Literature Review
by Elfadil Abdalla Mohameds, Nazar Zakis and Mohammad Marjans

Optimization of the Electrical Discharge Machining of Powdered Metallurgical High-Speed Steel Alloy using Genetic Algorithms
by Mohd Razif Idris and Imad Mokhtar Mosrati

Hospital Warehouse Management during the construction of a new building through Lean Techniques
by Carlotta Patrone, Maryna Mezzano Kozlova, Monica Brenta, Francesca Filauro, Donatella Campanella, Antonietta Ribatti, Elisabetta Scuderi, Tiziana Marini, Gabriele Galli and Roberto Revetria

Design and Optimization of ZnO Nanostructured SAW-Based Ethylene Gas Sensor with Modified Electrode Orientation
by Argha Sarkar, Padarthi Venkataramana, Nimmala Harathi, Thummuru Jyothsna and Neeruganti Vikram Teja

Prediction of Demersal Fishing Ground Associated with Coral Reefs in the Coastal Jepara Regency, Central Java, Indonesia Based on Sentinel 2a Imagery
by Kunarso, Muhammad Zainuri, Denny Nugroho Sugianto, Jarot Marwoto, Hariyadi and Muslim

Enabling 3D Heterogeneous Structures Towards Smart Chips: A Review
by Cheng Li, Mengfu Di, Zijin Pan and Albert Wang

Multiscale Texture Analysis and Color Coherence Vector Based Feature Descriptor for Multispectral Image Retrieval
by Devulapalli Sudheer and Rajakumar Krishnan

Optimization of Low Temperature Differential Stirling Engine Regenerator Design
by Hind El Hassani, Nour- Eddine Boutammachte and Sanae El Hassani

Angular Orientation of Steering Wheel for Differential Drive
by Rajesh Kannan Megalingam, Deepak Nagalla, Ravi Kiran Pasumarthi, Vamsi Gontu and Phanindra Kumar Allada

Electronic Steering of the Vehicle Axle with Force Feedback
by Tomas Pawlenka, Jaromir Skuta and Jiri Kulhanek

Controller Design Using Backstepping Algorithm for Fixed-Wing UAV with Thrust Vectoring System
by Shogo Hirano, Kenji Uchiyama and Kai Masuda

Implementation of Paraconsistent Logic Based PI Controller for TA Converter
by Tapas Kumar Mohapatra, Asim Kumar Dey and Krushna Keshab Mohapatra

An Integrated Framework for Pronominal Anaphora Resolution in Malayalam
by Ajees Arimbassery Pareed and Sumam Mary Idicula

Fuzzy Modelling using Firefly Algorithm for Phishing Detection
by Noor Syahirah Nordin, Mohd Arfian Ismail, Vitaliy Mezhuyev, Shahreen Kasim, Mohd Saberi Mohamad and Ashraf Osman Ibrahim

Mitigation Methods of Voltage Disturbances and Harmonic Currents of PWM Two-Level Inverter Supplying The 6-Pulse Diode Rectifier
by Jerzy Ryszard Szymanski and Marta Zurek-Mortka

Spectrum Requirements for Control and Non-payload Communication of Maritime Autonomous Surface Ship
by Ho Namgung

BPMN4 Collaboration: An Extension for collaborative Business Process
by Leila Amdah and Adil Anwar

The Impact of Manufactured Sand (M-Sand) as Partially and Fully Replacement of Fine Aggregate in Concrete
by Altamashuddinkhan Nadimalla, Siti Aliyyah Masjuki, Siti Asmahani Saad, Ean Lee Woen, Siti Maisarah Ali and Naseer Ulla

A Comparative Study of Safety Leading and Lagging Indicators Measuring Project Safety Performance
by Sevar Dilkhaz Salahaddin Neamat

Multimode Control and Simulation of 6-DOF Robotic Arm in ROS
by Rajesh Kannan Megalingam, Raviteja Geesala, Ruthvik Rangaiah Chanda and Nigam Katta

Effects of Split Ring Resonator (SRR) Metamaterial on the Radiation Pattern and Variation of the Heating Focus Point of the Microstrip Patch Antenna
by Mahsa Jahangiri and Saman Rajebi

Enhancing Bodily Movements of the Visually Impaired Children by Airflow
by Fang-Lin Chao, Hung-Chi Chu and Liza Lee

A Sustainable Multi-layered Open Data Processing Model for Agriculture: IoT Based Case Study Using Semantic Web for Hazelnut Fields
by ?ahin Aydin and Mehmet Nafiz Aydin

Fabrication of Glaze Material from Recycled Bottle Glass and Kaolin
by Agus Dwi Anggono, Elkana Bilak Lopo, Joko Sedyono and Tri Widodo Besar Riyadi

Stress Response Index for Traumatic Childhood Experience Based on the Fusion of Hypothalamus Pituitary Adrenocorticol and Autonomic Nervous System Biomarkers
by Noor Aimie Salleh, Malarvili Balakrishnan and Anna C. Whitttaker

Methodology for Assessing Synchronization Conditions in Telecommunication Devices
by Juliy Boiko, Ilya Pyatin, Oleksander Eromenko and Oleg Barabash

MIMO Fractional Order Control of a Water Tank Plant
by Arturo Rojas–Moreno and Juan Hernandez–Garagatti

ANN Based MRAC-PID Controller Implementation for a Furuta Pendulum System Stabilization
by Efrain Mendez, German Baltazar-Reyes, Israel Macias, Adriana Vargas-Martinez, Jorge de Jesus Lozoya-Santos, Ricardo Ramirez-Mendoza and Ruben Morales-Menendez and Arturo Molina

Performance of a Thermoelectric Powered by Solar Panel for a Large Cooler Box
by Mirmanto Mirmanto, Syahrul Syahrul, Made Wirawan, I Made Adi Sayoga, Agung Tri Wijayanta and Ilham Mahyudin
A Method for Mosaicking Aerial Images based on Flight Trajectory and the Calculation of Symmetric Transfer Error per Inlier
by Daniel Arteaga, Guillermo Kemper, Samuel G. Huaman Bustamante, Joel Telles, Leon Bendayan and Jose Sanjurjo

Extending Movable Surfaces with Touch Interaction Using the VirtualTablet: An Extended View
by Adrian H. Hoppe, Felix Marek, Florian van de Camp and Rainer Stiefelhagen

Skin Tissue Oxygen Saturation Prediction: A Comparison Study of Artificial Intelligence Techniques
by Audrey Huong and Xavier Ngu

Mutual Reduction in the Coupling of the MIMO Antenna Network Applied to the Broadband Transmission
by Chafai Abdelhamid and Hedi Sakli

Design of Efficient Convolutional Neural Module Based on An Improved Module
by Daihui Li, Chengxu Ma and Shangyou Zeng

Direction of Arrival Using 2-D Matrix Pencil Method
by Mohammed Amine ihedrane, Adiba EL FAdl and Seddik Bri

Determination of Submerged Breakwater Location for Coastal Protection in Panjang Island Waters Jepara
by Denny Nugroho Sugianto, Purwanto, Gentur Handoyo, Aulia Oktaviani, Kunarso, Muhammad Zainuri and Agus Indarjo

Project-Based Learning as a Tool of Enhancing of Entrepreneurial Attitude of Students
by Dana Paľová, Martin Vejačka and Lukáš Kakalejčík

Territorial Sales Redesign Using Geotechnical Tools
by Diana Sánchez-Partida, Moises Ernesto Sánchez-Castro, José-Luis Martínez-Flores, Patricia Cano-Olivos and Martin Straka

A Multi-Objective Voltage Optimization Technique in Distribution Feeders with High Photovoltaic Penetration
by Temitayo Olayemi Olowu, Mohamadsaleh Jafari and Arif Sarwat

Applications of Causal Modeling in Cybersecurity: An Exploratory Approach
by Suchitra Abel, Yenchih Tang, Jake Singh and Ethan Paek

A Critical Analysis of Usability and Learning Methods on an Augmented Reality Application for Zoology Education
by Diego Iquira-Becerra, Michael Flores-Conislla, Juan Deyby Carlos-Chullo, Briseida Sotelo-Castro, Claudia Payalich-Quispe and Carlo Corrales-Delgado

Novel Cost Function based Motion-planning Method for Robotic Manipulators
by Daniel Szabo and Emese Gincsaine Szadeczky-Kardoss
Racial Categorization Methods: A Survey
by Krina B. Gabani, Mayuri A. Mehta and Stephanie Noronha

Hybrid Autogyro: Model and Longitudinal Control for Wind Gust Energy Conversion Using Autorotation
by Jonathan Flores Santiago, Sergio Salazar Cruz and Rogelio Lozano Leal

Using Metaheuristics-Based Methods to Provide Sustainable Market Solutions, Suitable to Consumer Needs
by Ricardo Simões Santos, António João Pina da Costa Feliciano Abreu and Joaquim José Rodrigues Monteiro

Resource Selection Service Based on Neural Network in Fog Environment

by Nour Mostafa

A Big Data Security Layer Meta-Model Proposition
by Allae Erraissi and Abdessamad Belangour

A Comparative Analysis of ARIMA and Feed-Forward Neural Network Prognostic Model for Bull Services
by Jude B. Rola, Cherry Lyn C. Sta. Romana, Larmie S. Feliscuzo, Ivy Fe M. Lopez and Cherry N. Rola

Alleviation of Nonlinear Impact Using PAPR Hybrid Technique in CO-OFDM Systems
by Liqaa A. Al-Hashime, Sinan M. Abdul Satar, Ghaida A. Al-Suhail and Osama Saied

Linear Logic Synthesis of Multi-Valued Sequential Circuits
by Nikolay Butyrlagin, Nikolay Chernov, Nikolay Prokopenko and Vladislav Yugai

Pyrolysis of Beech Wood in A Continuous Drop Tube Reactor in Comparison to A Batch Reactor
by Chetna Mohabeer, Nourelhouda Boukaous, Antoinette Maarawi, Lokmane Abdelouahed, Abdeslam-Hassen Meniai, Mustapha Chikhi and Bechara Taouk

CSFs for the Implementation of the Hybrid Lean ERP System. Stakeholders Interactions
by Mariam Houti, Laila El Abbadi and Abdellah Abouabdellah

Promoting Continuous Professional Development Among Academics from A Vocational College by Using A Practical Workshop Based on Arduino Technology
by Arthur James Swart and Pierre Eduard Hertzog

The Usability Evaluation of Academic Progress Information System (SIsKA-NG)
by Gede Indrawan, I Made Agus Oka Gunawan and Sariyasa

Interaction Model and Respect of Rules to Enhance Collaborative Brainstorming Results
by Thierry Gidel, Andrea Tucker, Shigeru Fujita, Claude Moulin, Kenji Sugawara, Takuo Suganuma, Yuki Kaeri and Norio Shiratori

Smart Transmission Line Maintenance and Inspection using Mobile Robots
by Thongchai Disyadej, Surat Kwanmuang, Paisarn Muneesawang, Jatuporn Promjan and Kanyuta Poochinapan

Alternative Braking Method for Small Scaled-Wind Turbines Connected DC Green House with Analogical Experiment on Blade Destruction
by Anupa Koswatta, Faramarz Alsharif, Yasushi Shiroma, Mahdi Khosravy, Shiro Tamaki and Junji Tamura

Mechanical Behaviors of Kaolin Powder Filler Polypropylene/Low Density Polyethylene Blends
by Pham Thi Hong Nga

A Simple Modelling Tool for Fast Combined Simulation of Interconnections, Inter-Symbol Interference and Equalization in High-Speed Serial Interfaces for Chip-to-Chip Communications

by Davide Menin, Thomas Bernardi, Alessio Cortiula, Martino Dazzi, Alessio De Prà, Mattia Marcon, Marco Scapol, Andrea Bandiziol, Francesco Brandonisio, Andrea Cristofoli, Werner Grollitsch, Roberto Nonis and Pierpaolo Palestri

Understanding Human Perception of Vibrotactile Feedback in Walking and Running Tasks
by Emel Demircan, Elliot Recinos, I-Hung Khoo, Sharon Teng and Will Wu

Securing Hybrid SDN-based Geographic Routing Protocol using a Distributed Trust Model
by Lylia Alouache, Mohamed Maachaoui and Rachid Chelouah

Determinants of Students' Actual use of the Learning Management System (LMS): An Empirical Analysis of a Research Model
by Charles Buabeng-Andoh and Charles Baah

Commercialization Process of Disruptive Innovations in Corporate Ventures and Spinoff Companies: A Comparison
by Javier Nieto Cubero, Saheed Adebayo Gbadegeshin and Carolina Consolación Segura

Activerest: Design of A Graphical Interface for the Remote use of Continuous and Holistic Care Providers
by Nuno Martins, Jéssica Campos and Ricardo Simoes

Classification of Timber Load on Trucks
by Jan Sikora and David Fojtík

ROS Based Multimode Control of Wheeled Robot
by Rajesh Kannan Megalingam, Santosh Tantravahi, Hemanth Sai Surya Kumar Tammana, Nagasai Thokala, Hari Sudarshan Rahul Puram and Naveen Samudrala

Neural Network-based Efficient Measurement Method on Upside Down Orientation of a Digital Document
by Yeji Shin, Youngone Cho, Hyun Wook Kang, Jin-Gu Kang and Jin-Woo Jung

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 201-211 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

An Algorithm for Automatic Measurement of KI-67 Proliferation Index in Digital Images of Breast Tissue

Guillermo Kemper[*,1], David Atencia[1], Ivan Ortega[1], Roberto Kemper[2], Alejandro Yabar[3]

[1]*Faculty of Engineering and Architecture, School of Electronic Engineering, University of San Martín de Porres, Lima 12, Peru.*

[2]*Institute of Pathology and Molecular Biology Arias Stella, Lima 11, Peru.*

[3]*Department of Anatomy Pathology, National Hospital Edgardo Rebagliati Martins, Lima 11, Peru.*

ARTICLE INFO

Article history:
Received: 17 October, 2019
Accepted: 24 December, 2019
Online: 22 January, 2020

Keywords:
Breast Cancer
Image Processing,
Immunohistochemistry
KI-67

ABSTRACT

This paper proposes an algorithm aimed at quantifying the expression of KI-67 protein in digital images of breast biopsy tissue samples obtained through an optical microscope. The algorithm allows to obtain a report on the quantity of non-proliferating and proliferating cells through the detection and quantification of KI-67. The sample analysis via software aims to reduce the level of subjectivity in the diagnosis of diseases such as breast cancer. The algorithm proposed involves the application of statistical image processing techniques, adaptive thresholds, object segmentation and color filtering. A method of analysis and quantification of overlapping cells is also proposed to improve the efficiency of the algorithm. The results of the method proposed were satisfactory, as they were highly correlated with those obtained by visual inspection by pathologists.

1. Introduction

Cancer is a public health problem worldwide, as evidenced by its high rates of incidence and mortality. In Peru, for example, the Lima Cancer Registry reported, in 2011 alone, about 34,000 new cases of cancer [1]. Likewise, according to the Peruvian Epidemiological Surveillance System, in the period 2006-2011, breast cancer was the second most common among Peruvian women [2], in order to choose the correct therapeutic treatment for breast cancer, guidelines often combine conventional predictors to estimate relapse/mortality risk. [3][4]

In medical research, some biomarkers have been validated for the detection and estimation of the disease progression, including breast cancer. Among them are the estrogen receptor (ER), the progesterone receptor (PR), HER2 and KI-67 [3][4]. Through a quantitative analysis, KI-67 provides a reliable estimate of the proliferative activity in a tissue. Its application to biopsies in patients with breast cancer gives doctors some evidence to treat the disease [4][5].

KI-67 is a protein that is present in all cell cycle phases (G1, S, G2, M), except for the resting phase (G0); its presence is, therefore, quite intense and large in neoplastic proliferations [3]. For this reason, the quantification of KI-67-positive cells provides valuable information to determine the magnitude of the proliferative activity.

To prove the existence of protein KI-67 in a breast biopsy tissue sample, a specific anti-protein is used with a chromogen additive allowing visualization, denominated MIB-1, this cell staining is made in a process called Immunohistochemistry (IHC). Finally, the specialist analyses the sample (in a subjective way) by eyeballing inspection through a microscope to give a report to the patient on the proliferative activity of KI-67-positive cells. According to the St. Gallen Consensus (S. Bustreo, S. Osella-Abate, P. Cassoni et al, 2016), the KI-67 index is one of the most important prognostic markers used by oncologists to determine the treatment of ER-positive breast cancer patients [6]. The number of proliferating cells in proportion to the total number of cells in a particular tissue area, called hereinafter 'proliferative index' (PI), is divided into 2 levels of prediction for Ki-67: Luminal A, if the proliferative index is less than 20%; and Luminal B, if it is over or equivalent to 20% [6].

However, subjective evaluation can lead to errors in diagnosis due to several factors: room lightning, stress, fatigue, concentration, visual ability, experience, etc. They often cause that the results provided by pathologists are uncorrelated and have high

[*]Corresponding Author: Guillermo Kemper, Email: gkemperv@usmp.pe

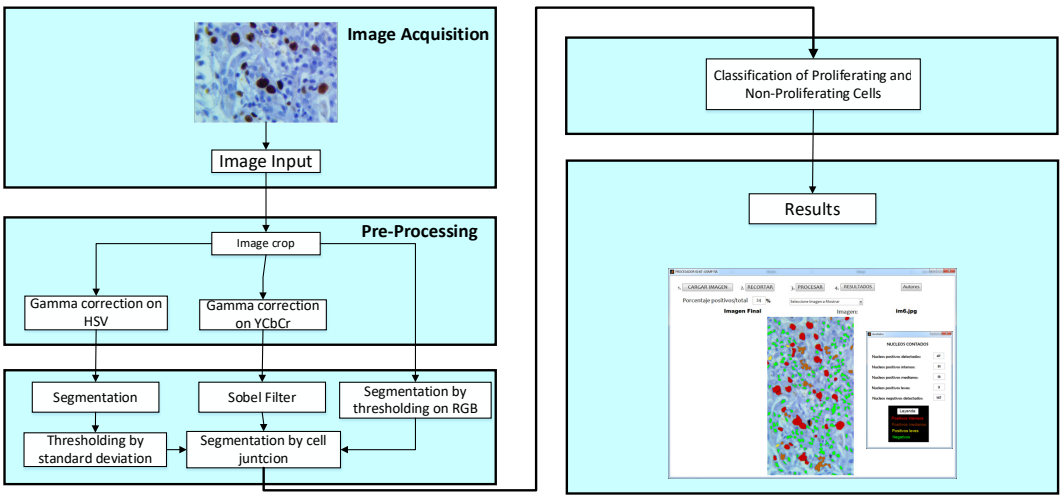


Fig. 1. Block diagram of the algorithm proposed.

variance [7][8], another factor that influences is the heterogenity of the stained sample [9]. Possible solutions to this issue have been proposed in previous research, but the main problems affecting cell detection in similar software remain: imperfection in collecting data such as noise, colour distortion and deformity in histopathological material at the stage of sample pigmentation, therefore a good calibration of the staining machine between IHC test should be made to avoid difference in pigmentation, this create a inter-laboratory variability on identical set of tumors [10].

In [11], for example, the results of the quantification algorithm were successful in analysing KI-67 with different image processing techniques such as thresholding, segmentation and labelling. However, the use of the ultra-erosion technique damaged valuable information for the evaluation of the sample.

In [12], they implement an algorithm based on mathematical description of cell morphology added with Support Vector Machine (SVM), obtaining great accordance between the digital image analysis and the "eyeballing" method.

On the other hand, M. Bouzid et al. show in [13] an interesting proposal with the use of special hardware: a multispectral camera with programmable light source. However, it increased the cost and availability of the solution. F. Kabir and N. Yusoff used in [14] the method of Random Forest (RF), consisting of predictive tree-like structures, each independent of the other, that obtain their data from different random vectors, but same distribution. RF has proved to perform better than other machine learning techniques, such as neural networks, support vector machines and k-nearest neighbors. The results of the tests, however, showed an accuracy of 72% in differentiating between malignant and benign tissue.

In this paper, an alternative solution is proposed on the basis of a digital image processing algorithm designed to quantify and provide more objective results regarding the level of cell proliferation based on marker KI-67. The algorithm involves statistical image processing techniques [15], adaptive thresholds [15][16], object segmentation and colour filtering [17]. Similarly, it proposes an overlapping cell analysis and quantification method to improve the efficiency of cell detection. These techniques have been used together to obtain satisfactory results based on the problem. The results were indeed satisfactory, as presented at the end of this document.

In the following sections, the processing stages of the algorithm proposed in this paper will be described.

2. Description of the proposed algorithm

The proposed algorithm involves different processing parts and stages, as shown in Fig. 1. The scheme proposed contains the stages of image acquisition, pre-processing, segmentation, classification and counting. These steps will be described below.

2.1. Image acquisition

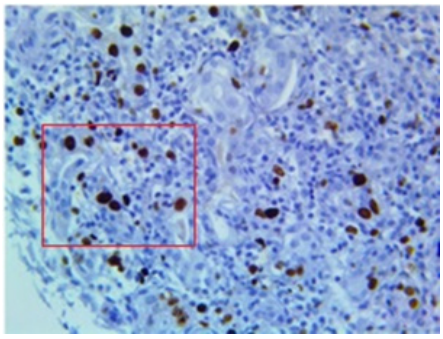
Before being digitalised, the biopsy sample goes through an IHC process, which mixes the laminated sample with chemicals and special antibodies to achieve proper homogenous staining.
Currently, the procedure is performed automatically using the Ventana BenchMark XT stainer in line with the producer's manual. After staining, the sample is placed under an Olympus BX53 digital microscope with a 40x magnifying lens, 50 images were chosen to study, all were given by chief of the department of anatomy pathology of Hospital Edgardo Rebagliati Martins, all samples were chosen randomly. The images of the biopsy sample (pre-treated by IHC) are obtained from a digital camera and an optical microscope. In this case, camera resolution was $M0$=2748 rows by $N0$=3584 columns with 40x zoom. The colour format of the original image (Fig. 2a) is RGB (Red, Green and Blue) called $I(x, y, z)$ with 12 bits per primary colour component (36 bits/pixel). The primary components can be expressed as $F_R(x, y), F_G(x, y)$ and $F_B(x, y)$.

All the parameters used in this paper were chosen taking in account the procedure mentioned previously to convert the biopsy sample into a digital image.

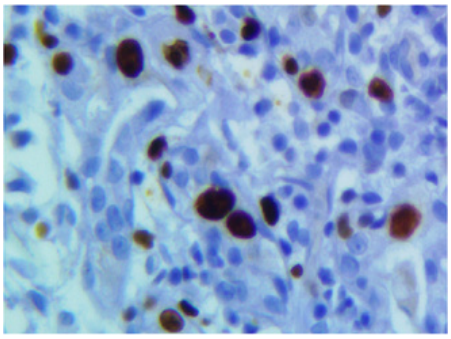
2.1.1. Choosing the area of interest

The image can be cropped to select an area of interest. This procedure grants the pathologist the freedom to choose the area to analyse according to their own criteria and preference.
The selected area will have a resolution of M rows by N columns (this values may vary depending on the area selected by the

specialist). The format is also RGB with primary components $\boldsymbol{I_R(x,y)}, \boldsymbol{I_G(x,y)}$ and $\boldsymbol{I_B(x,y)}$.
Figs. 2a and 2b show the original image and the cropped image, respectively.



(a)



(b)

Fig. 2. (a) Original RGB image where the red square is the area to be zoomed. (b) Zoomed cropped image.

2.2. Pre-Processing

Pre-processing aims to improve the quality of the image obtained in order increase the efficiency of segmentation algorithms. This process involves the following image processing blocks: Gamma correction on HSV (Hue, Saturation and Value) and Gamma correction on YCbCr (Luminance, Chrominance Blue, Chrominance Red).

2.2.1 Gamma correction on HSV

The cropped image expressed in its primary components $I_R(x,y), I_G(x,y)$ e $I_B(x,y)$ is converted to the HSV component $I_V(x,y)$ [15].

Gamma correction is applied with logarithmical effect to this component so as to enhance the contrast between the background (epithelium) and the proliferating cells. This will further increase the efficiency of segmentation algorithms. The conversion applied can be expressed as [17]:

$$I_V(x,y) = max\big(I_R(x,y), I_G(x,y), I_B(x,y)\big) \quad (1)$$

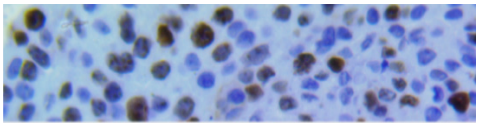
Where $max\big(a_1, a_2, a_3\big)$ function generates the maximum value among a_1, a_2 and a_3.

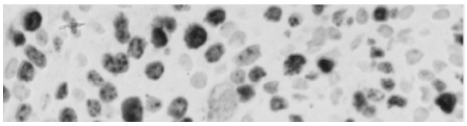
The Gamma correction applied to $I_V(x,y)$can be expressed as [13]:

$$I_{Vg}(x,y) = round\left(\frac{I_V(x,y)}{255}\right)^{\gamma} \quad (2)$$

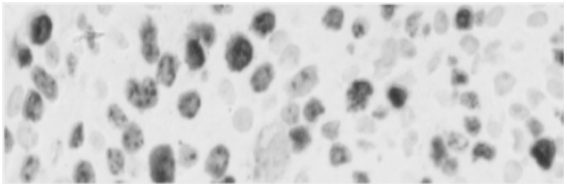
where γ = 0.9 and round(x) return the approximate value to the nearest integer.

Fig. 3a shows image $I(x,y)$ in RGB format; Fig. 3b shows $I_V(x,y)$ after applying (1) to $I(x,y)$; and Fig. 3c shows $I_{Vg}(x,y)$, where non-proliferating or blue cells obtain a colour scheme similar to that of the background, making proliferating cells stand out. Likewise, in Fig. 3c, partially proliferating cells become evident.





(b)



(c)

Fig. 3. Images: (a) $I(x,y)$, (b) $I_V(x,y)$ and (c) $I_{Vg}(x,y)$

2.2.2. Gamma correction on YCbCr

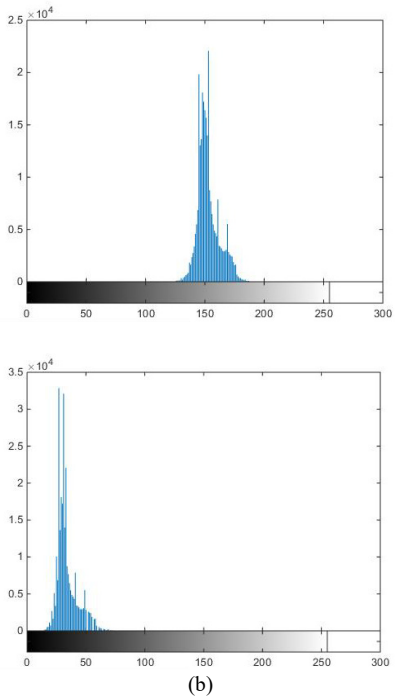
In parallel with the process above, the cropped image expressed in its primary components $I_R(x,y), I_G(x,y)$ and $I_B(x,y)$ is converted to the YCbCr component $I_{Cb}(x,y)$ the conversion can be expressed as [15]:

$$I_{Cb}(x,y) = 128 - 37.797 I_R(x,y) - 74.203\, I_G(x,y) + 112\, I_B(x,y) \quad (3)$$

Gamma correction is applied to this component with γ = 4 (exponential effect). This allows to displace Cb colour schemes from non-proliferating cells. Displacement from the original Cb dimensional space, shown in Fig. 4a. is shown in Fig. 4b. The image resulting from the process is expressed as $I_{Cb2}(x,y)$, as shown in Fig. 5.
The correction applied allows to separate colour schemes and improve future segmentation of the two types of cells.

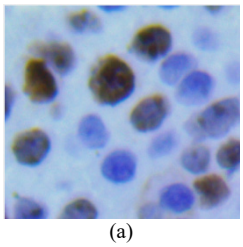
2.3. Segmentation

In this process, objects corresponding to proliferating and non-proliferating cells are segmented. Segmentation is made from the corrected images $I_{Vg}(x,y)$ and $I_{Cb2}(x,y)$ obtained from the original image (Fig. 6).

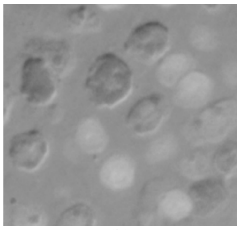


(b)

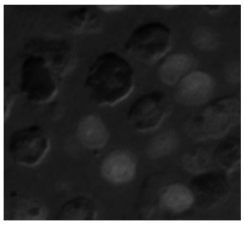
Fig. 4. (a) Original Cb (histogram) dimensional space. (b) Displaced Cb2 (histogram) dimensional space.



(a)



(b)



(c)

Fig. 5. (a) Original image RGB. (b) Image $I_{Cb}(x,y)$. (c) Image $I_{Cb2}(x,y)$ after application of Gamma correction.

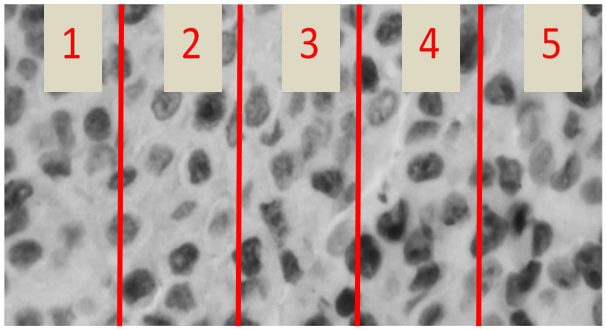


Fig. 6. Image $I_{Vg}(x,y)$ to be subjected to the proliferating cell segmentation process

2.3.1. Segmentation of proliferating cells

To separate the background (cell tissue of no interest for diagnosis) from the proliferating cells, it was decided to use local adaptive thresholds, which result from the probability density function (obtained by histogram) and first order statistical measures (mean and variance). The steps to achieve this process are as follows:

Step 1:

Image $I_{Vg}(x,y)$ is segmented in L' segments. Image segments will have a spatial resolution of M rows by $\frac{N}{L'}$ columns. The segments of image $I_i(x,y)$ can be expressed as:

$$I_i(x,y) = I_{Vg}\left(x, y + i\left(\frac{N}{L'}\right)\right) \tag{4}$$

Where $x = 0,1,\ldots,M-1$; $y = 0,1,\ldots,\frac{N}{L'}-1$ and $i = 0,1,\ldots,L'-1$.

In this case, we have $L' = 5$, because a more even segment is obtained in terms of noise and illumination (Fig.6.).

Step 2:
The histogram of each i segment is determined, which is defined as $h_i(r)$. In this case r represents each colour scheme from 0 to 255 (Fig. 7).

Step 3:
The mean value of each segment i is obtained:

$$\bar{r}_i = \frac{L'}{M.N}\sum_{r=0}^{255} r h_i(r) \tag{5}$$

Step 4:
The standard deviation for each segment i is obtained:

$$\sigma_i = \sqrt{\frac{L'}{MN}\sum_{r=0}^{255}(r-\bar{r}_i)^2\, h_i(r)} \tag{6}$$

Step 5:
The threshold u_i to be used locally for each segment in order to separate cell tissue from the proliferating cell area is the result of the subtraction of the mean and standard deviation:

$$u_i = \bar{r}_i - \sigma_i \tag{7}$$

Step 6:
Thresholding is then applied to the segment:

$$\hat{I}_i(x,y) = \begin{cases} 1\,; & I_i(x,y) \geq u_i \\ 0 & ; \quad I_i(x,y) < u_i \end{cases} \tag{8}$$

Step 7:
Finally, each segment thresholded $\hat{I}_i(x,y)$ is placed in the original order so as to reconstruct an image with spatial resolution of M

rows and N columns, which will be called $I_{pro}(x,y)$ and will contain the proliferating cell area as shown in Fig. 8.

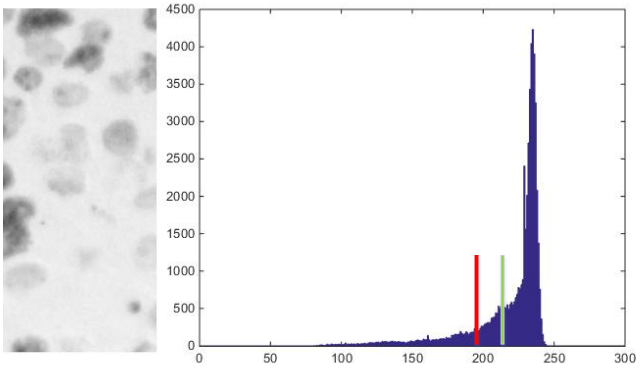


Fig. 7. Example: Segment 1 of $I_{Vg}(x,y)$) and its histogram (in this case, the green bar indicates the mean value of the histogram and the red bar the threshold obtained).

2.3.2. Segmentation by Sobel filter and morphology

After identifying proliferating cells, it is necessary to spot non-proliferating cells in order to find the percentage of cells with proliferative activity in the image under evaluation $I(x,y)$.
For this process, image $I_{Cb2}(x,y)$ obtained in the step above is used. The procedure is detailed below:

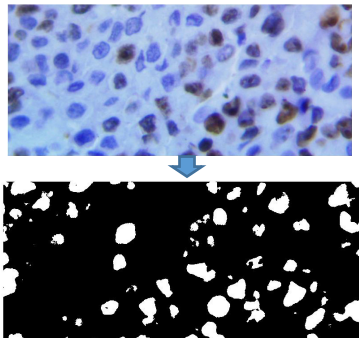


Fig. 8. Original Image and resulting binary image $I_{pro}(x,y)$ after the proliferating cell segmentation process.

Step 1:
A thresholding filter is applied to $I_{Cb2}(x,y)$ in order to segment non-proliferating cells. Thresholding can be expressed as:

$$I_{Cb3}(x,y) = \begin{cases} 0 \quad ; \quad I_{Cb2}(x,y) < 41 \\ I_{Cb2}(x,y) \quad ; \quad other\ way \end{cases} \tag{9}$$

The threshold was obtained from the study and evaluation of image histograms $I_{Cb2}(x,y)$ corresponding to various samples obtained. The result of the process is shown in Fig. 9.

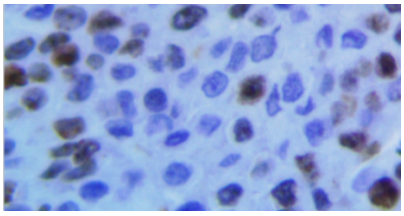
Step 2:
Then, the Sobel filter [17] for edge detection is applied to binary images.
In this process, $I_{Cb3}(x,y)$ is convolved with each of the Sobel masks (horizontal and vertical) which are defined in matrix as [14]:

$$\mathbf{M_v} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{10}$$

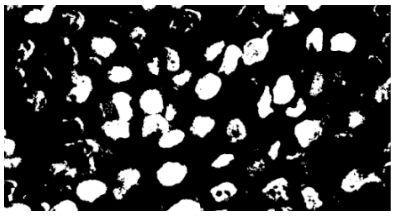
$$\mathbf{M_h} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{11}$$

The resulting image can be expressed as (binary):

$$I_{Sobel}(x,y) = \begin{cases} 1 \ ; \ I'_{Sobel}(x,y) > 6 \\ 0 \quad ; \quad other\ way \end{cases} \tag{12}$$



(a)



(b)

Fig. 9. (a) Original image. (b) Resulting image $I_{Cb3}(x,y)$

Where:

$$I'_{Sobel}(x,y) = |I_{Cb3}(x,y) * M_h(x,y)| + |I_{Cb3}(x,y) * M_v(x,y)| \tag{13}$$

Step 3:
$I_{Sobel}(x,y)$ is subjected to a "closure" morphological operation which can be expressed through the following matrix equation:

$$\mathbf{I}_{clo} = ([(\mathbf{I}_{Sobel} \oplus \mathbf{B}_v) \oplus \mathbf{B}_h] \ominus \mathbf{B}_v) \ominus \mathbf{B}_h \tag{14}$$

Where $\oplus$ indicates the dilation operation and $\ominus$ the erosion operation.
The structural elements used can be expressed as:

$$\mathbf{B_h} = [\mathbf{1\ 1\ 1}] \tag{15}$$

$$\mathbf{B}_v = \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \\ \mathbf{1} \end{bmatrix} \tag{16}$$

In order to reform the cells detected and not to distort the visual result, gap-filling is applied to the matrix cells $\mathbf{I}_{clo}$. This operation can be expressed as [16]:

$$\mathbf{I}_{clo}^{k} = \left(\mathbf{I}_{clo}^{k-1} \oplus \mathbf{D}_p\right) \cap \dot{\mathbf{I}}_{clo}$$
$$k = 1,2,3,\ldots,K0 \tag{17}$$

Where k indicates the number of filling iteration, $\cap$ denotes the intersection operation, and $\dot{\mathbf{I}}_{clo}$ expresses the complement of matrix $\mathbf{I}_{clo}$. Filling is made through $K0$ iterations until condition $\mathbf{I}_{clo}^{k} = \mathbf{I}_{clo}^{k-1}$ is satisfied. The structuring element $\mathbf{D}_p$ used can be expressed as:

$$\mathbf{D}_p = \begin{bmatrix} \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \end{bmatrix} \tag{18}$$

Finally, the result of this stage is show on Fig.10 and placed in matrix $\mathbf{I}_{Sobelf}$:

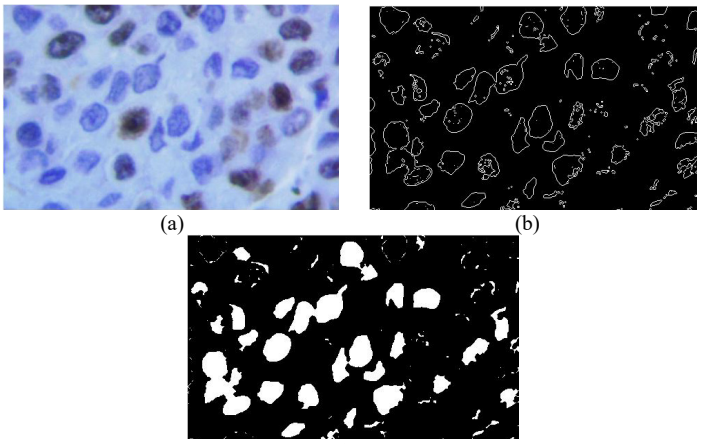
$$\mathbf{I}_{Sobelf} = \mathbf{I}_{clo}^{K0} \tag{19}$$

2.3.3. Segmentation by thresholding on primary components

At this stage, thresholding is applied to the primary RGB components to separate proliferating and non-proliferating cells from the epithelium (background). The thresholds used were obtained from experimental tests performed with images from different tissue samples. The procedure helps remove the area of the epithelium and highlight cells of interest. The steps of the procedure are as follows:

Step 1:
The primary components of the original image **I** are thresholded (Fig 11a):

$$Z_R(x,y) = \begin{cases} 0\ , & I_R(x,y) < 128 \\ 1\ , & I_R(x,y) \geq 128 \ \vee \ I_G(x,y) \geq 128 \end{cases} \tag{20}$$



(a) (b) (c)

Fig. 10. (a) Cropped original image. (b)Resulting Image after applying Sobel Filter. (c)Final Image $I_{Sobel}(x,y)$ after gap-filling.

$$Z_G(x,y) = \begin{cases} 0\ , & I_G(x,y) < 128 \\ 1\ , & other\ way \end{cases} \tag{21}$$

$$Z_B(x,y) = \begin{cases} 0\ , & I_B(x,y) < 128 \\ 1\ , & I_B(x,y) \geq 128 \ \vee \ I_G(x,y) \geq 128 \end{cases} \tag{22}$$

Step 2:
A matrix of zeros $I_{cel}(x,y)$ with spatial resolution $M \times N$ is generated. Then some pixels of $I_{cel}(x,y)$ are changed to 1 from the following condition:

$$I_{cel}(x,y) = \begin{cases} 1\ , & (Z_R(x,y) = 1\ ) \wedge \\ & (\ Z_G(x,y) = 1)\ \wedge \\ & (Z_B(x,y) = 1) \\ 0\ , & other\ way \end{cases} \tag{23}$$

$I_{cel}(x,y)$ finally constitutes a template where cell pixels have value 1 (Fig. 11b).

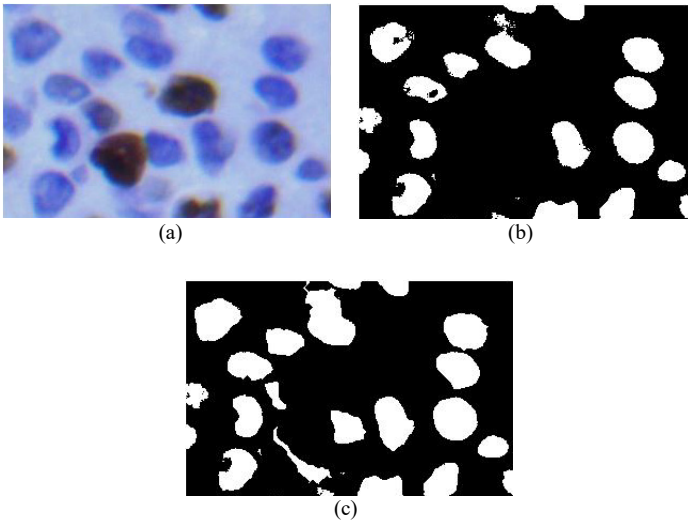
Step 3:
Finally, the matrix of non-proliferating cells (Fig. 11c) is obtained from the following expression:

$$I_{nopro}(x,y) = \begin{cases} 1\ , & (I_{cel}(x,y) = 1\ ) \vee \\ & (\ I_{sobelf}(x,y) = 1\ ) \wedge \\ & (I_{Pro}(x,y) \neq 1) \\ 0\ , & other\ way \end{cases} \tag{24}$$

2.3.4. Segmentation by cell junction

The combination of the proliferating and non-proliferating cell areas results in a binary image containing two types of segmented cells. This image is obtained as follows:

$$I_{total}(x,y) = \begin{cases} 1\ , & \left(I_{pro}(x,y) = 1\right) \vee \\ & \left(\ I_{nopro}(x,y) = 1\right) \\ 0\ , & other\ way \end{cases} \tag{25}$$



(a) (b) (c)

Fig. 11. (a) Cropped original image. (b) Image $I_{cel}(x,y)$. (c) Image $I_{nopro}(x,y)$.

However, one of the problems still present when the cells detected are counted is that some objects are made up of two or more joint or overlapped cells, which results in the algorithm giving a wrong diagnosis. In that sense, as in an acquisition with 40x magnification the average size of a cell is between 200 and 1500 pixels, we can consider that an object of $I_{total}(x,y)$ with a size above 1500 pixels is composed of more than one cell. This will definitely affect the counting, as an object is in principle composed of a single cell. To solve the problem, the following cell separation algorithm is proposed in order to minimise the counting error:

Step 1:
Consider $I_j(x_j, y_j)$ a binary segment of $I_{total}(x,y)$ that contains object j with more than 1500 pixels. In this case, $x_j = 0,1,2 \ldots, m_j - 1 \ldots$, and $y_j = 1,2, \ldots, n_j - 1$; i.e. the segment size is $m_j \times n_j$ pixels.

Step 2:
We perform image rotation for different angles θ ($\theta_0 = 0°, \theta_1 = 15°, \theta_2 = 30°, \theta_3 = 60°, \theta_4 = 75°$ and $\theta_5 = 90°$). In this case, the rotation axis passes through the centre of image $I_j(x_j, y_j)$. The range of 15° between rotation angles was chosen based on the best results obtained in various tests of cell separation. Each turn generates an image that can be expressed as $I_j^{\theta}(x'_j, y'_j)$,where θ indicates the rotation angle. The spatial dimensions of each $I_j^{\theta}\left(x'_j, y'_j\right)$ is expressed as m_j^{θ} rows by n_j^{θ}columns.

Step 3:
The horizontal projection $I_j^{\theta}\left(x'_j, y'_j\right)$ is obtained:

$$P_j^{\theta}(y'_j) = \sum_{x'j=0}^{m_j^{\theta}-1} I_j^{\theta}\left(x'_j, y'_j\right) \tag{26}$$

Step 4:
It is important to assess the central area of protection for each $P_j^{\theta}(y'_j)$, which is why the furthest areas of the vector are deleted:

$$\hat{P}_j^{\theta}\left(y'_j\right) = \begin{cases} P_j^{\theta}\left(y'_j\right) , & \left(y'_j > q_0.n_j^{\theta}\right) \vee \\ & \left(y'_j < q_1.n_j^{\theta}\right) \\ 0 , & other\ way \end{cases} \tag{27}$$

Where $q_0 = 0.3$ and $q_1 = 0.7$.

Step 5:
Value $\hat{P}_j^{\theta}(y_0(\theta, j))$is obtained, which is defined as the minimum value of the projection $\hat{P}_j^{\theta}\left(y'_j\right)$. In this case, $y_0(\theta, j)$ is the position where $\hat{P}_j^{\theta}\left(y'_j\right)$ is the minimum value.

Step 6:
The $\mathbf{W}_j$ vector is then obtained:

$$\mathbf{W}_j = \begin{bmatrix} \hat{P}_j^{\theta_0}(y_0(\theta_0, j)) \\ \hat{P}_j^{\theta_1}(y_0(\theta_1, j)) \\ \vdots \\ \hat{P}_j^{\theta_5}(y_0(\theta_5, j)) \end{bmatrix} \tag{28}$$

Step 7:
The minimum value of vector $\mathbf{W}_j$ is obtained, which is defined as $\hat{P}_j^{\theta_{L0}}(y_0(\theta_{L0}, j))$, where θ_{L0} is the angle corresponding to the minimum value of $\mathbf{W}_j$.

Step 8:
Then $Z'_j(x'_j, y'_j)$ is obtained from $I_j^{\theta_{L0}}(x'_j, y'_j)$:

$$Z'_j\left(x'_j, y'_j\right) = \begin{cases} 0 \quad ; \quad y'_j = y_0(\theta_{L0}, j) \\ I_j^{\theta_{L0}}\left(x'_j, y'_j\right) \quad ; \quad other\ way \end{cases} \tag{29}$$

Step 9:
Rotation of $-\theta_{L0}$ over $I_j^{\theta_{L0}}\left(x'_j, y'_j\right)$ is applied to restore the original orientation (although distortion is present due to the interpolation applied to the rotation process). After rotation, image $V_j(x'_j, y'_j)$ is obtained.

Step 10:
Finally, the block of image $V_j(x'_j, y'_j)$ is inserted in $I_{total}(x,y)$ (in the region where $I_j(x_j, y_j)$ was extracted).

$$I_{total}(xa_j + x'_j, ya_j + y'_j) = V_j(x'_j, y'_j) \tag{30}$$

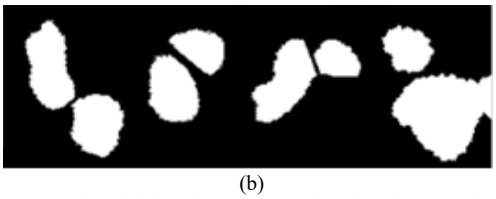
Where xa_j and ya_j are the coordinates of the first pixel segment extracted $I_j(x_j, y_j)$.
Fig. 12a shows original areas of $I_{total}(x,y)$ containing objects made up of more than one cell. Fig. 12b shows the results obtained after applying the algorithm of cell separation.

2.4. Classification and counting of proliferating and non-proliferating cells

From the total number of cells detected in image $I_{total}(x,y)$, we proceed to classify and count proliferating and non-proliferating cells.



(a)



(b)

Fig. 12. (a) Samples with junction. (b) Samples after the separation process.

Cooperation with pathologists was managed so that the parameters of cell classification were correct and a good diagnosis could be achieved. In this case, doctors requested to classify proliferating cells into three sub groups: strongly proliferating, moderately proliferating and poorly proliferating. The classification criterion is based on the size of the cell area affected by the KI-67 protein.

Image $I_{total}(x, y)$is subjected to a labelling algorithm [16] that will label each cell detected as a k object. In this case, k is the object counter detected in the image. For a total of Q objects, we have $k = 0,1,2, \ldots Q-1$.

2.4.1. Classification of Proliferating and Non-Proliferating Cells

For the classification stage, counters $\mathrm{Total_{pro}} = 0$ (proliferating cell counter) and counter $\mathrm{Total_{npro}} = 0$ (non-proliferating cell counter) are initialised.

$\mathrm{T_{obj}}$ is defined as the total number of objects or cells contained in $\mathrm{I_{total}(x, y)}$. In this context, k is taken to be the object or cell number to be processed. In this case $\mathrm{k} = 0,1, \ldots, \mathrm{T_{obj}} - 1$. We make $\mathrm{k} = 0$ and carry out the following procedure:

Step 1:
A segment $I_k(x^*, y^*)$ of image $I_{total}(x, y)$ is extracted consisting of M_k rows and N_k columns, where $x^* = 0,1, \ldots, M_k - 1$ and $y^* = 0,1, \ldots, N_k - 1$. In this case, $I_k(x^*, y^*)$ is the image segment containing the object k detected in $I_{total}(x, y)$.

Step 2:
A flag f_k is updated through the following expression:

$$f_k = \sum_{x^*=0}^{M_k-1} \sum_{y^*=0}^{N_k-1} I_k(x^*, y^*). I_{pro}\left(x_{0_k} + x^*, y_{0_k} + y^*\right) \quad (31)$$

Where x_{0_k} and y_{0_k} are the coordinates of the first pixel of $I_k(x^*, y^*)$ in $I_{total}(x, y)$.

Step 3:
If $f_k \neq 0$, the cell under evaluation is considered proliferating and the counter $Total_{pro}$ increases by one unit. Otherwise, the counter $Total_{npro}$ increases. Within the requirements of doctors, a sub classification of proliferating cells was made in 3 levels of proliferation: strongly proliferating, moderately proliferating and poorly proliferating.
In this way, when $f_k \neq 0$, steps 4 to 9 are executed.

Step 4:
A temporary matrix $C_k(x^*, y^*)$ consisting of M_k rows and N_k columns is generated and applied:

$$C_k(x^*, y^*) = \begin{cases} 1\,; \; si\; I_{pro}\left(x_{0_k} + x^*, y_{0_k} + y^*\right) \neq 0 \\ 0\,; \qquad other\; way \end{cases} \quad (32)$$

Step 5:
We count each pixel belonging to an area that reacted to KI-67:

$$Apro_k = \sum_{x^*=0}^{M_k-1} \sum_{y^*=0}^{N_k-1} C_k(x^*, y^*) \quad (33)$$

Step 6:
Similarly, for the non-proliferating cell area, a matrix $C'_k(x^*, y^*)$ consisting of $M_k$rows and N_k columns is generated and applied:

$$C'_k(x^*, y^*) = \begin{cases} 1\,, I_{nopro}\left(x_{0_k} + x^*, y_{0_k} + y^*\right) \neq 0 \\ 0\,, \qquad other\; way \end{cases} \quad (34)$$

Step 7:
The pixels of the regions that did not react to the protein KI-67 are counted:

$$Anpro_k = \sum_{x^*=0}^{M_k-1} \sum_{y^*=0}^{N_k-1} C'_k(x^*, y^*) \quad (35)$$

Step 8:
Then the intensity factor of each cell is found:

$$R_k = \frac{Apro_k}{Anpro_k + Apro_k} \quad (36)$$

Step 9:
Flags of proliferation fp_k, fm_k and fd_k are updated:

$$fp_k = \begin{cases} 1\,, & R_k \geq 0.7 \\ 0\,, & other\; way \end{cases} \quad (37)$$

$$fm_k = \begin{cases} 1\,, & R_k > 0.2 \wedge R_k < 0.7 \\ 0\,, & other\; way \end{cases} \quad (38)$$

$$fd_k = \begin{cases} 1\,, & R_k \leq 0.2 \\ 0\,, & other\; way \end{cases} \quad (39)$$

For each k where $f_k \neq 0$ there is only one active flag (fp_k, fm_k and fd_k).
Three counters (initialised to zero), defined as strong, medium and weak, are updated. They are updated as follows: if $fp_k = 1$, the strong counter is incremented by one; if $fm_k = 1$, the medium counter is increased by one; if $fd_k = 1$, the weak counter is increased by one. These counters store the amount of subclassified cells. Note that strong + medium + weak = $Total_{pro}$. This data is granted to the pathologist for evaluation and diagnosis, as shown in Table 1.
Then, k increases by one unit and we go back to step 1. The procedure ends when $k = T_{obj}$.

Step 10:
After evaluating all the cells, we find the percentage of proliferating cells and obtain:

$$\%Diagnosis = \left(\frac{Total_{pro}}{Total_{pro} + Total_{npro}}\right) \times 100 \quad (40)$$

Table 1. Results after test.

Total counted nucleus	Quantity	Percentage (%)
Negative nucleus	94	88
Positive nucleus	13	12
Total	107	100
Positive counted nucleus		
Strong positive nucleus	2	15
Medium positive nucleus	7	54
Weak Positive Nucleus	4	31

3. Results

To prove that the software can be used by specialists, the results obtained by it had to be validated against the results provided by doctors; they were requested to make the same analysis on equal samples. The doctors that carried out the tests are specialised in pathology and work in different institutions in the city of Lima.

Table 2, 3, 4, 5 and 6 show the results obtained in different cases. When comparing the results obtained by the application against the diagnosis given by doctors, it was concluded that, in samples with a clear KI-67-active cell rate (either in a very small or very high percentage in proportion to the total number of cells), doctors using the eyeballing technique obtain results very similar to those of the application. But when samples present some difficulty at the time of counting, it is observed that the results among doctors are poorly correlated with each other and only a few can give a diagnosis similar to that of the application. This is because doctors are subject to different stimuli that influence their performance and they have different counting methods, such as hotspots, random areas, or groups near adjacent hotspots [7], while the software will always proceed in the same way.

In Figure 13 it is shown a section of an evaluated image by both the medical specialist and the developed algorithm. Five cells have been enumerated and marked in a rectangle, named: cell 1, cell 2, cell 3, cell 4 and cell 5.

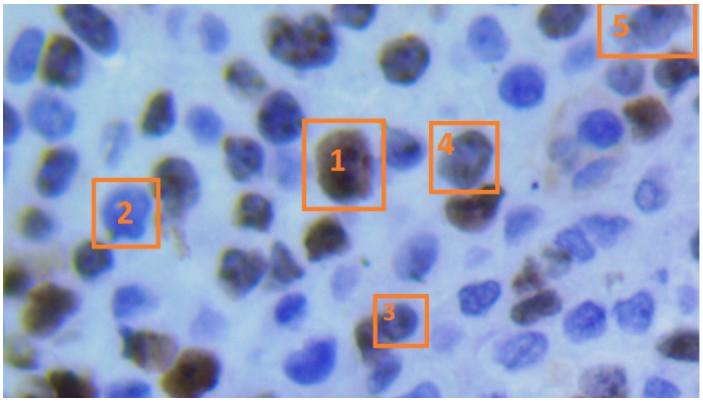


Fig. 13. Cells evaluated by specialist and the proposed algorithm.

For cell 1, both medical specialist and proposed algorithm confirmed that represent a proliferative cell, meaning it was stained in brown color by the IHC process.
For cell 2, again both specialist and algorithm concurred that the cell was non-proliferative, stained with a blue tone by the IHC process. In case of cell 3 and 5, the staining was lower than cell 1, specialist marked the cell as non-proliferative and the algorithm marked as proliferative, is in this scenario in which the subjectivity of human eye turn the proliferation index and can occur divergence in the diagnosis. In case of cell 4, shows partial staining and both specialist and algorithm matched as proliferative.

To demonstrate the error trend, the table below shows the calculus of coincidence between the software and the specialists, and its measuring using the Kappa coefficient.

Defining Kappa coefficient as [18]:

$$\tilde{K} = \frac{\bar{T} - \bar{T}_e}{1 - \bar{T}_e} \tag{41}$$

Define $\bar{T}$ as:

$$\bar{T} = \frac{1}{N_0}\sum_{i=1}^{N_0} \beta_i \tag{42}$$

Where N_0 is the total number of category, in this case two categories $N_0 = 2$: Luminal A, and Luminal B, as explained on [6], Luminal A if diagnosis was <20% and Luminal B if was ≥20%.

β_i is defined as:

$$\beta_i = \frac{1}{\alpha(\alpha - 1)}\left[\left(\sum_{j=1}^{N_0} \alpha_{ij}^2\right) - \alpha\right] \tag{43}$$

Where α is the number of occurrences from a category, α_{ij} is the number of times the specialist chose a diagnosis in the category i and the software chose category j.

Finally:

$$\bar{T}_e = \sum_{j=1}^{N_0} \varepsilon_j^2 \tag{44}$$

Where ε_j is calculated using the following expression:

$$\varepsilon_j = \frac{1}{N_0 \alpha}\left(\sum_{i=1}^{N_0} \alpha_{ij}\right) \tag{45}$$

To calculate precision and recall of specialist vs software, $F - score$ value was calculated [19]:

$$F - score = \left(\frac{(\rho^2 + 1).precision \times recall}{\rho^2.precision \times recall}\right) \tag{46}$$

Where precision and recall were calculated as:

$$precision = \frac{wp}{wp + lp} \tag{47}$$

And:

$$recall = \frac{wp}{wp + ln} \tag{48}$$

Where:
wp: true positive, a cell that specialist said was proliferative and algorithm said was proliferative.
lp: false positive, a cell that specialist said wasn't proliferative and algorithm said it was proliferative.
ln: false negative, a cell that said specialist said was proliferative but algorithm said wasn't proliferative.

$F - score$ is considered balanced so:

$$\rho = 1 \tag{49}$$

Finally, $F - score$ is calculated as:

$$F - score = \frac{2wp}{(2wp + lp + ln)} \tag{50}$$

The average result for F-score of the samples was of 0.8764, showing a high concordance between the software and specialist. The software detects a greater number of cells than the specialist, so the F-score is not perfect, however, these cells not accounted were mostly discarded by specialist errors.
As shown in the Tables 2, 3, 4, 5 and 6 there is a remarkable variation among specialists in the Luminal B diagnosis made by counting. The poor agreement in the diagnosis is due to the difference between stained and non-stained nuclei as rapidly observed by the specialist. Nevertheless, two specialists gave the same result as the software; it should be noted that a proliferation index >20% is needed to share the same diagnosis.
Tables 7, 8 and 9 show the difference among specialists when giving their diagnosis. The Kappa index among specialists is not constant, an average of 0.6863, which demonstrates, once again, the difference generated by the subjective analysis by eyeballing [9][10].

Table 2. Comparison Software vs. Specialist 1

Specialist 1: Kappa $\tilde{K} = 0.689$			
Software \ Specialist 1	Luminal A	Luminal B	TOTAL
Luminal A	7	3	10
Luminal B	0	9	9
TOTAL	7	12	19

Table 3. Comparison Software vs. Specialist 2

Specialist 2: Kappa $\tilde{K} = 0.689$			
Software \ Specialist 2	Luminal A	Luminal B	TOTAL
Luminal A	7	3	10
Luminal B	0	9	9
TOTAL	7	12	19

Table 4. Comparison Software vs. Specialist 3

Specialist 3: Kappa $\tilde{K} = 1$			
Software \ Specialist 3	Luminal A	Luminal B	TOTAL
Luminal A	10	0	10
Luminal B	0	9	9
TOTAL	0	9	19

Table 5. Comparison Software vs. Specialist 4

Specialist 4: Kappa $\tilde{K} = 1$			
Software \ Specialist 4	Luminal A	Luminal B	TOTAL
Luminal A	10	0	10
Luminal B	0	9	9
TOTAL	0	9	19

Table 6. Comparison Software vs. Specialist 5

Specialist 5: Kappa $\tilde{K} = 0.787$			
Software \ Specialist 5	Luminal A	Luminal B	TOTAL
Luminal A	10	0	10
Luminal B	2	7	9
TOTAL	12	7	19

Table 7. Comparison Specialist 1 vs. Specialist 2

Specialist 1 vs Specialist 2: Kappa $\tilde{K} = 0.774$			
Specialist 2 \ Specialist 1	Luminal A	Luminal B	TOTAL
Luminal A	6	1	7
Luminal B	1	11	12
TOTAL	7	12	19

Table 8. Comparison Specialist 2 vs. Specialist 4

Specialist 2 vs Specialist 4: Kappa $\tilde{K} = 0.596$			
Specialist 4 \ Specialist 2	Luminal A	Luminal B	TOTAL
Luminal A	7	0	7
Luminal B	4	8	12
TOTAL	11	8	19

Table 9. Comparison Specialist 1 vs. Specialist 5

Specialist 1 vs Specialist 5: Kappa $\tilde{K} = 0.689$			
Specialist 5 \ Specialist 1	Luminal A	Luminal B	TOTAL
Luminal A	7	3	10
Luminal B	0	9	9
TOTAL	7	12	19

4. Conclusions

In this publication, an algorithm to quantify the proliferative index of the KI-67 marker in histological samples of breast tissue has been developed that proves to give results more reliable than results given by specialist, thanks to image processing techniques, this algorithm can help to provide a better diagnostic to treat breast cancer when applied the result with other biomarkers to diagnose cancer.

A software application was developed. It meets the standards and criteria that pathologists use for the evaluation of the immunohistochemistry method and it is approved by specialists, as they consider that, when testing the software in their work centres, results are reliable.

The iterative algorithm of cell separation is computationally heavy and must be optimised for the comfort of physicians.

In developing the software, in cooperation with medical advisors, new limitations and standards emerged; this is the only way to consolidate a multidisciplinary project.

The quality of the images provided must have some standard sharpness and correct pigmentation to be processed by the application.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] E. Poquioma. “Situación del cáncer de mama en el Perú”. Presentation. Instituto Nacional de Enfermedades Neoplásicas (INEN), 2008.

[2] C. Ramos, D. Venegas. “Análisis de la Situación del Cáncer en el Perú”. Presentation. Dirección General de Epidemiología, Ministerio de Salud del Perú, 2013.

[3] S. Coronato, G. Laguens, O. Spinelli, W. Di Girolamo. “Marcadores tumorales en cancer de mama” MEDICINA (Buenos Aires), 63, 73-82, 2002.

[4] B. Hirata, B. Karina, et al. "Molecular markers for breast cancer: prediction on tumor behavior." Disease markers, 2014. http://dx.doi.org/10.1155/2014/513158

[5] H. Xin et al. “Ki-67 is a valuable prognostic predictor of lymphoma but its utility varies in lymphoma subtypes: evidence from a systematic meta-analysis.” BMC cancer, 14, 153, 2014. http://doi:10.1186/1471-2407-14-153

[6] S. Bustreo, S. Osella-Abate, et al. “Optimal Ki67 cut-off for luminal breast cancer prognostic evaluation: a large case series study with a long-term follow-up” Breast Cancer Res Treat, 157(2), 363–371, 2016. 10.1007/s10549-016-3817-9

[7] L. Fulawka, A. Halon. “Ki-67 evaluation in breast cancer: The daily diagnostic practice” Indian J. Pathol Microbiol, 60(2), 177-184, 2017. http://doi:10.4103/IJPM.IJPM_732_15

[8] Gudlaugsson, E., Skaland, I., Janssen, E. A., Smaaland, R., Shao, Z., Malpica, A., ... & Baak, J. P. (2012). Comparison of the effect of different techniques for measurement of Ki67 proliferation on reproducibility and prognosis prediction accuracy in breast cancer. Histopathology, 61(6), 1134-1144.

[9] Zhong, F., Bi, R., Yu, B., Yang, F., Yang, W., & Shui, R. (2016). A comparison of visual assessment and automated digital image analysis of Ki67 labeling index in breast cancer. PloS one, 11(2), e0150505.

[10] Koopman, T., Buikema, H. J., Hollema, H., de Bock, G. H., & van der Vegt, B. (2018). Digital image analysis of Ki67 proliferation index in breast cancer using virtual dual staining on whole tissue sections: clinical validation and inter-platform agreement. Breast cancer research and treatment, 169(1), 33-42.

[11] V. Anari, P. Mahzouni, R. Amirfattahi “Computer-aided Detection of Proliferative Cells and Mitosis Index in Inmunohistichemically Images of Meningioma” In 6th Iranian Conference on Machine Vision and Image Processing, Isfahan, Iran. IEEE, 2010. 10.1109/IranianMVIP.2010.5941151

[12] Grala, B., Markiewicz, T., Kozłowski, W., Osowski, S., Słodkowska, J., & Papierz, W. (2009). New automated image analysis method for the assessment of Ki-67 labeling index in meningiomas. Folia Histochemica et Cytobiologica, 47(4), 587-592.

[13] M. Bouzid, A. Khalfallah, A. Bouchot, M. Selim, F. Marzani. “Automatic Cell Nuclei Detection: a Protocol to Acquire Multispectral Images and to compare results between colour and multispectral images” in SPIE Photonics West, BIOS, Imaging, Manipulation and Analysis of Biomolecules, Celles and Tissues XI, San Francisco, United States, 2013. doi.10.1117/12.2001980

[14] K. Ahmad, N. Yusoff. “Breast Cancer Types Based on Fine Needle Aspiration Biopsy –Data Using Random Forest Classifier” in I 13th International Conference on Intellient Systems Design and Applications, Bangi, Malaysia. IEEE, 2013. doi.10.1109/ISDA.2013.6920720

[15] R. Gonzalez, R. Woods. Digital Image Processing, 3rd. Ed., Prentice Hall, 2008.

[16] T. Chun-Ming, L. His-Jian. “Binarization of Colour Document Images via Luminance and Saturation Colour Features” IEEE Transactions on Image Processing, 11(4), 2002. 10.1109/TIP.2002.999677

[17] R. Gonzales, R. Woods, E. Eddins. Digital Imaging Processing Using Matlab, 2nd Ed., Mcgraw Hill, 2010.

[18] J.L Fleiss. “Measuring Nominal Scale Agreement Among Many Raters” Psycological Bulletin, 76(5): 378-382, 1971. http://dx.doi.org/10.1037/h0031619

[19] M. Sokolova, N. Japkowicz, S. Szpakowicz. “Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation” In AI 2006: Advances in Artificial Intelligence, Lecture Notes in Computer Science (vol. 4304), 2006. DOI: 10.1007/11941439_114

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 01-07 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Single-leg Standing Ability and Lower Limb Movement Analysis of Collegiate Footballers and Sedentary Students

Jiun Sien Lau, Vina Phei Sean Tan, Mohd Akmal Vahidi, Siti Hanisah Linggi Mohd Azam, Mohd Hafezi Mat Zain, Azilawati Azizan, Mazra Othman, Rosniwati Ghafar*

Exercise and Sports Science, School of Health Sciences, Universiti Sains Malaysia, Health Campus, 16150, Malaysia

ARTICLE INFO

Article history:
Received: 30 September, 2019
Accepted: 10 December, 2019
Online: 15 January, 2020

Keywords:
Balancing
Lower limbs
Footballer
Sedentary

ABSTRACT

Stability or balance is an integral component to perform daily activities without incurring injury or to be dependent on others. Sportsmen tend to have better balance than non-sports people but less is known about the single-leg balance ability. Furthermore, few studies analyzed the dynamic phase of single-leg stance that may contribute to better overall balance. Sports like football tend to have instances where the player's non-dominant leg keeps them in an upright position while the dominant leg kicks, passes and stops the ball. We aim to study the single-leg balance between collegiate footballers and sedentary students in eyes-closed (EC) and eyes-opened (EO) conditions and their contributing components to keep a body in an up-right position. Twenty collegiate footballers and 20 sedentary students conducted the unipedal stance test (UPST) on each leg with EO and EC conditions while standing on a force platform. We captured center of pressure (CoP) distance travelled, stance duration and using a 3D motion capture system, we assessed lower limb movement at six different anatomical sites. Results showed that footballers had better overall balance compared to sedentary students only in the non-dominant leg, EC condition with 12 footballers versus four sedentary students completing the full 45s stance ($p=0.01$) The other three UPST conditions did not show significant differences between groups. The CoP distance in the initial dynamic state and total UPST were both significantly shorter in footballers than sedentary students ($p<0.05$) during the non-dominant leg, EC stance. Our multivariable linear regression model significantly predicted time for UPST on non-dominant leg with EC up to 76.8% ($p<0.001$) with the first 5-s of greater trochanter movement significantly contributing to total time taken for UPST in footballers. Overall, playing football may enhance balance control intrinsically especially for the non-dominant side while being less reliant on visual input to maintain balance.

1. Introduction

Balance is a crucial motor skill stemming from muscle synergies that lessens centre of pressure (CoP) displacement to effectively keep our body upright in a standing position, during locomotion and body orientation [1]. These muscle synergies are coordinated by the central nervous system with various inputs from our visual, vestibular and somatosensory systems [2]. In sports like gymnastics, balance is key to excel in the sport, however, other sports like football and basketball, static and dynamic balance are just as important although less apparent in their movements [3].

In football (considered as soccer in North America), most players shoot, pass and stop the ball with their dominant leg while the non-dominant leg supports their body weight [4]. Footballers often play and train on various turf conditions and wear cleated or non-cleated shoes. These conditions challenge trained footballers to develop better balance while mastering their football skills. Hence, top performing footballers have better posture stability and use less visual information to maintain their balance [5].

It is important to assess balance ability as it may relate to the level of sports performance and injury prevention. Single-limb and two-limb stance tests with eyes opened and with eyes closed are commonly used to test standing balance [6]. The unipedal stance

*Corresponding Author: Rosniwati Ghafar, rosnikk@usm.my

test (UPST) is a common balance and standing ability assessment that can be easily performed and does not require additional equipment [7]. However, UPST lacks normative values for different age groups and balance ability changes with age [8]. Yet, UPST could detect changes of postural balance after 12 weeks of recreational soccer training and running in untrained men [9]. Researchers concluded that soccer training and high-intensity interval running enhanced the balance ability in their study participants [9]. As the UPST mainly measures the time taken to perform the test, we can also assess the center of pressure (CoP) during UPST using advanced technology to further understand the mechanisms of balance.

There are various types of force or pressure assessment systems to quantify CoP. Force platforms detect the ground reaction force and CoP. CoP can also be assessed from the force distribution across a sensor grid by other pressure assessment systems [10].

In sports biomechanics, motion data are used to study and observe human performance. The three-dimensional (3-D) motion capture system consists of infrared cameras with reflective markers attached to anatomical bony landmarks [11]. This system captures and measures motion that was hard to assess previously. 3-D motion capture plays an important role in sports to analyze physical limitations and for movement optimization [11]. Through the analyses, details about injury mechanism and movement will lead to new preventive methods and be used to improve a player's technique and performance [12].

There are two distinct phases in UPST based on Jonsson and colleagues' research [13]. The first phase is the first 5 s of UPST and also known as the dynamic phase where a rapid reduction in force variability can be seen. The better control a person has in the dynamic phase, the better their balance will be to complete the 45-s UPST. The second phase is known as the static phase where the individual maintains a certain level of force variability that enables balance to be achieved [13]. Jonsson and colleagues concluded that the dynamic phase was key in assessing UPST [13]. Yet, we know little about lower limb influence on single-leg stance in young adults [14].

To our knowledge, studies that compared the balance ability between collegiate footballers and sedentary students are limited despite the possibility of recreational football play could enhance balance ability in previously untrained men [9]. Furthermore, balance is an important functional fitness component and should be cultivated to avoid balance-related injuries [15]. Therefore, our study aims to investigate if indeed simple football play could be related to better balance in young adults. We will compare the standing ability between collegiate footballers and sedentary students. We will also conduct further analysis on the dynamic phase of UPST to better understand lower limb control during the single-leg balance test.

2. Methodology

2.1. Participants

We recruited male collegiate students aged between 18 to 25 years old and grouped them as footballers (n= 20) or sedentary students (n=20) based on our study's participant criteria. Footballers need to have a minimum of 2 years of football competition experience at collegiate level and/or exercise at least 3 times or more a week. Footballers need not be part of any specific training other than their recreational football games during non-competition time. Sedentary students were those who exercised less than twice a week and do not participate in any sport competition. We excluded individuals with health conditions that may affect balance such as diabetes mellitus, musculoskeletal disorders, vestibular impairments and those with any lower limb injuries in the past 6 months. All participants provided informed consent and the Human Research Ethical Committee of Universiti Sains Malaysia provided ethical approval for this study.

2.2. Study Procedure

We measured participants' height (cm) and weight (kg) using a stadiometer and digital weight scale, respectively. Measures were done in duplicates and a third measure was obtained if difference between two measures were 0.4 cm in height or 0.2 kg in weight. We used the averaged value of measures obtained and calculated body mass index (BMI, kg/m^2) of all participants.

Participants conducted the UPST for both dominant and non-dominant legs with i) eyes-opened (EO) and ii) eyes-closed (EC) conditions. Participants were instructed to kick a ball and the leg used was considered the dominant leg. Test was terminated early if participants i) used arms to balance their body, ii) touched the floor with the raised foot, iii) moved the standing foot to maintain their body posture, such as rotating the foot on the force platform or iv) opened their eyes in the EC condition. [8]. Each participant conducted three UPST trials and had 30 s of rest between trials. We recorded the total time taken to maintain the single-leg stance and stopped the UPST after 45 s.

Participants stood on a force platform (Bertec, USA) and perform UPST by raising one leg as shown in Figure 1. The sampling rate of the platform was set at 120 Hz. Participants stood barefooted on the level platform with the foot positioning marked on the force platform. This ensures uniformity and standardization of feet placement on the force platform which minimise variations between trials [16].

In the attempts to maintain posture in the UPST, body muscles are constantly adjusting and this creates 'sway' or body movement that can be captured by the force platform [17]. The force platform captures CoP in the x- and y-axis in a constant flow and consequently, we assessed the distance travelled by CoP. Lesser distance and smaller CoP range usually indicates better balance skills [18].

A 3D motion capture system (Qualysis AB, Gothenburg Sweden) recorded lower limb movement and time during UPST. Six motion analysis cameras were placed surrounding the force platform as shown in Figure 2. In order to record movement of the limb during UPST, we attached 12 reflective markers to specific anatomical bony landmarks on the lower extremities. The reflective markers were attached to the greater trochanter, lateral side of the knee condyle, ankle and heel, and at the first and fifth metatarsal of both legs. The motion camera system captured the movement made during different UPST conditions and CoP from the force platform. Total distance was calculated from the movement of COP during performing UPST. The average time of the three trials and coordinates of CoP were extracted [8].

The first five seconds of UPST were further analysed to understand more about the dynamic phase of UPST. We used the results of the best trial out of the three attempts in this analysis. The best trial was selected based on the duration of the UPST and the shortest distance of CoP travelled during the same UPST trial.

2.3. Statistical Analysis

Statistical analyses were conducted using SPSS version 24.0 with significance set at p-value <0.05. Time taken to complete UPST had a skewed distribution and we used Mann-Whitney to compare the differences between groups. We used independent t-test to compare the differences between footballers and sedentary students for the normally distributed data from different UPST conditions.

To understand the relation of lower limb movement on UPST, we conducted Pearson's correlation tests to identify the related anatomical sites during the dynamic phase that contributes to the time taken to complete UPST. Subsequently, we included relevant variables in a multivariable linear regression analysis.



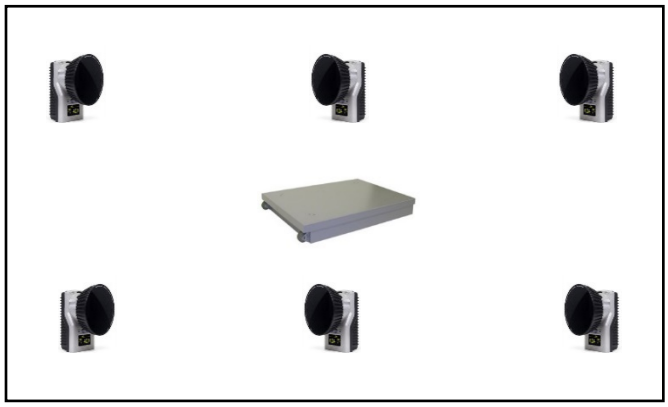


Figure 2: The field set-up of infrared cameras relative to the force platform

3. Results

Footballers were significantly younger than sedentary students by about 14 months but were similar in build (height, weight and BMI) as displayed in Table 1.

Table 1. Study participant characteristics in mean±standard deviation.

	Football (n=20)	Sedentary (n=20)	*p*-value
Age (yr)†	21.4 ± 1.1	22.6± 1.2	0.002
Height (m)	1.69±0.07	1.66±0.05	0.089
Weight (kg)	66.3±14.0	60.7±14.0	0.056
BMI (kg/m^2)	23.2±4.2	22.3±5.8	0.099

†*Differences between groups analysed using independent t-test while other variables were analysed using Mann-Whitney test.*
BMI, body mass index.

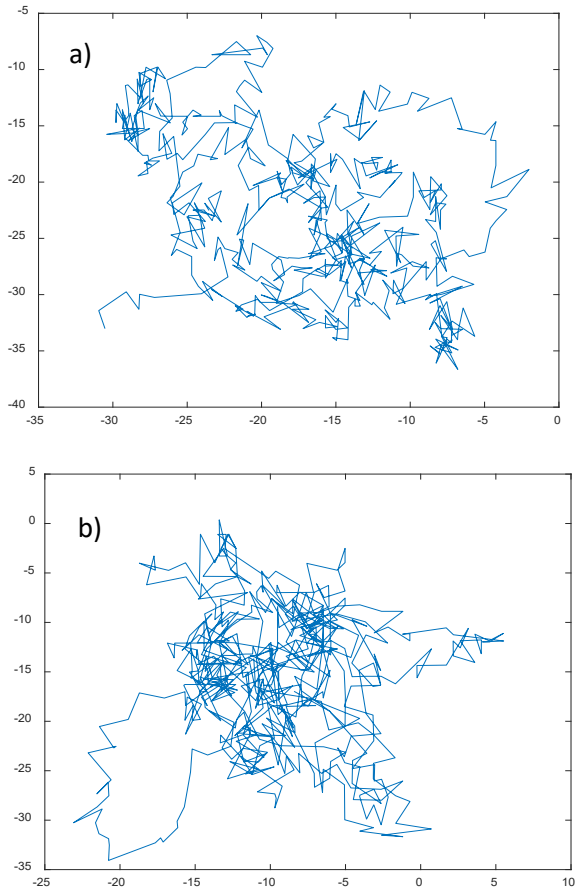


Figure 3: Examples of the 'best' or smallest range of CoP travel pathway for a) footballer and b) sedentary student during the non-dominant, eyes-closed condition.

3.1. Time Taken to Perform UPST

All footballers had no issues completing the UPST at the maximal 45 s when eyes were opened while standing on their dominant foot and only one person did not make the 45 s mark when standing on their non-dominant leg with EO (Table 2). Sedentary students did not perform any different than footballers on their dominant leg with EO or EC. However, there were three times more footballers (n=12) that could complete UPST till 45 s compared to sedentary students (n=4) when performing on their non-dominant leg with EC (p=0.01).

As data was not normally distributed in the time taken to perform UPST in four different conditions, we did not detect any significant differences (p>0.05) in time between groups as tested using the Mann-Whitney test.

3.2. Distance Travelled of CoP

Footballers were no different in their total CoP distance travelled during UPST than sedentary students using their dominant leg in both EO and EC conditions. From Table 2, total CoP distance on non-dominant leg, EO UPST were no different between groups too ($p>0.05$). In non-dominant leg, EC UPST, footballers significantly swayed less compared to sedentary students ($p=0.01$).

3.3. Distance Travelled of CoP in Dynamic Phase

Examples of CoP patterns of a footballer versus a sedentary student are presented in Figure 3 and 4 during the dynamic phase or the first 5-s portion of UPST. We selected the best (smallest) and worst (largest) range of CoP for both groups respectively in the non-dominant leg, EC session. Footballer CoPs showed less complexity of pattern within the range compared to a sedentary students' CoP travel path.

The 5-s CoP distance travelled were thus significantly different ($p=0.021$) between footballers and sedentary students when analysed using Mann-Whitney test (Table 2). The rest of the conditions were no different in 5-s CoP distance reported between footballers and sedentary students.

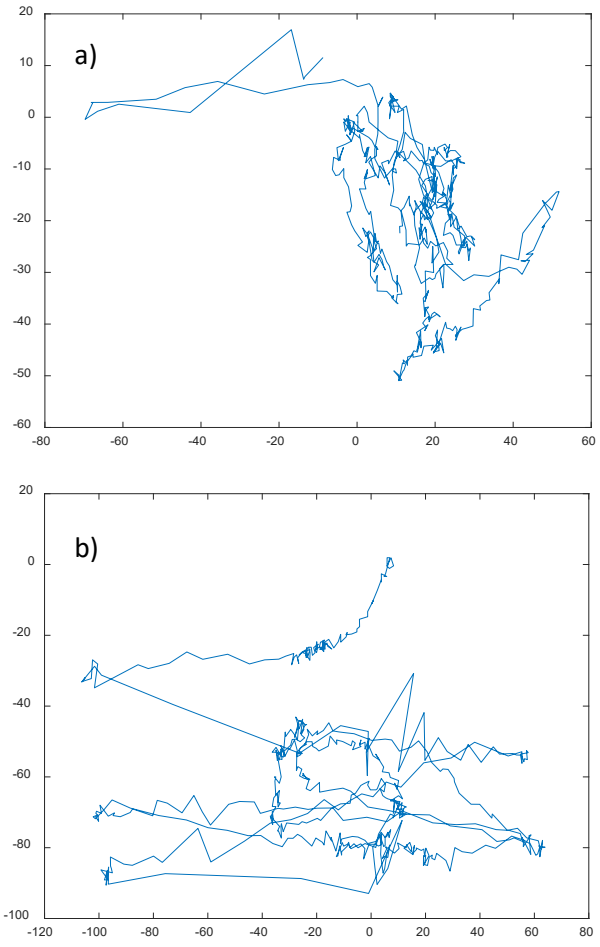


Figure 4: Example of a 'worst' or largest range of CoP travel pathway for a) footballer and b) sedentary student during the non-dominant, eyes-closed condition.

3.4. Non-dominant Leg, EC UPST Time and Lower Limb Control During Dynamic Phase

In Table 3, we used the best trial to further study the contribution of lower limb movement to overall time taken to complete UPST. Footballers could maintain their UPST significantly longer than sedentary students using the non-dominant leg in the EC condition.

We analysed the first 5-s of lower limb function on single-leg balance during the dynamic phase toward overall time taken to complete UPST. Interestingly, the lower limb movement in all six anatomical sites were not significantly different between footballers and sedentary students (Table 3). Yet, we observed significant negative correlation of all lower limb movement sites to time taken to complete UPST (Table 4). However, we had to remove four data sets due to noise in lower limb movement analysis.

Subsequently, we included the lower limb movements in the first 5-s of UPST to predict time to complete UPST. We analysed footballers and sedentary students separately. We found that different sites of lower limb movement were not predictive of overall time to complete UPST in sedentary students ($p>0.05$). However, our regression model in footballers (Table 5) showed that the overall model significantly predicted up to 76.1% ($p<0.001$) of total time to complete UPST. In that regression model, although we have ankle, knee and greater trochanter as predictor variable, the movement from the greater trochanter was the significant variable to predict time to complete UPST. The first and fifth metatarsal and heel sites were removed due to high correlation within the regression model (variance inflation factor >5).

4. Discussion

The objective of the study is to compare the standing balance ability between male collegiate football players and sedentary students and to understand the influence in the dynamic phase toward overall single-leg standing. We found that footballers on non-dominant leg with eyes-closed (EC) could significantly balance better than sedentary students as they completed more trials, had a longer balance time and CoP distance was shorter in the dynamic and static phase. Furthermore, we found that the footballers' greater trochanter control in the early dynamic phase significantly predicts overall balance performance in the non-dominant leg, EC condition. Whereas lower limb movements of sedentary students may not be helpful in maintaining balance during an EC, single-leg stand.

Differences in balance ability between sports/trained with untrained individuals are known. Specifically, in football, two other studies have showed that the highly skilled footballers relied less on visual information to maintain a stable posture [5,19]. Naturally, footballers were also known to be proficient in using their non-dominant leg as well as their dominant side [20]. Studies of other sports that included gymnastics and swimming also showed better postural control and efficient use of somatosensory and otolithic input to maintain body posture [21, 22]. Results of sports contributing to better balance is even apparent in recreational players whereby simple football training resulted in adapted visual, somatosensory and vestibular senses that improved

Table 2. Outcomes of number of individuals that completed UPST till 45 s and the average of three trials for i) time taken to complete unipedal stance test (UPST), ii) total CoP distance in complete UPST and iii) CoP distance in the first 5-sec of UPST in footballers and sedentary students, reported as median (inter-quartile range).

	Footballer (n=20)				Sedentary students (n=20)			
	Dominant leg		Non-dominant leg		Dominant leg		Non-dominant leg	
	EO	EC	EO	EC	EO	EC	EO	EC
Completed max. 45-s UPST (n)	20	11	19	12	19	9	18	4*
Duration of UPST (s)	45.0 (0.0)	23.0 (32.6)	45.0 (0.0)	31.2 (27.1)	45.0 (0.0)	20.4 (21.6)	45.0 (13.7)	24.4 (21.8)
Total CoP distance (mm)†	144.8 (2.6)	140.1 (3.9)	144.3 (3.0)	140.5 (3.9)	145.8 (0.6)	140.8 (5.0)	145.8 (0.6)	144.1 (4.7)*
CoP distance in first 5-s (mm)	133.8 (0.1)	134.0 (0.7)	133.8 (0.1)	134.3 (0.5)	133.8 (0.1)	134.1 (0.3)	133.8 (0.1)	134.0 (0.3)*

EO, eyes-opened; EC, eyes-closed.
**Significant differences between footballers and sedentary students in similar leg and eye conditions with p-value <0.05 as analysed by Mann-Whitney test or independent t-test.*
† Differences between groups were analysed using independent t-test and results reported as mean and standard deviation.

Table 3. Comparison of time, centre of pressure (CoP) range of x- and y-axis and six anatomical marker movements (mm) in non-dominant, eyes-closed condition of unipedal stance test (UPST) best trial† between footballers and sedentary students.

	Median (IQR)		
	Football (n=20)	Sedentary (n=20)	*p*-value
Complete UPST			
Time (s)	45.0 (21.2)	29.4 (29.2)	0.044*
CoP x-axis range (mm)	45.4 (19.0)	41.3 (20.2)	0.646
CoP y-axis range (mm)	41.4 (6.4)	42.1 (12.9)	0.626
Distance travelled in first 5-s of UPST			
1st metatarsal (mm)	100.6 (73.4)	103.5 (65.9)	0.534
5th metatarsal (mm)	73.6 (40.1)	72.1 (54.6)	0.787
Ankle (mm)	64.5 (44.8)	61.5 (30.4)	0.607
Heel (mm)	43.4 (26.3)	55.8 (18.7)	0.162
Knee (mm)	146.1 (107.7)	142.2 (87.5)	0.978
Greater trochanter (mm)	118.0 (68.7)	134.4 (54.6)	0.555

†Best trial is defined as the longest time taken and with the shortest CoP distance travelled during UPST.
IQR, inter-quartile range.

Table 4: Correlation between time for complete UPST time (s) and lower limb maker distances (mm) during the dynamic phase (n=36).

Lower limb sites	Time	
	r	*p*-value
First metatarsal	-0.50	0.001
Fifth metatarsal	-0.46	0.003
Ankle	-0.38	0.012
Knee	-0.44	0.004
Heel	-0.55	<0.001
Greater trochanter	-0.61	<0.001

Table 5: Multivariable linear regression model to predict the influence of greater trochanter, knee and ankle on time taken to complete unipedal stance test in collegiate footballers (n=18).

Variables	β	B	SE	*p*	95% CI
Greater trochanter	-0.77	-0.22	0.06	0.002	-0.35, -0.09
Knee	-0.09	-0.02	0.04	0.644	-0.09, 0.06
Ankle	-0.06	-0.02	0.05	0.685	-0.13, 0.09
Constant	-	69.1	5.2	<0.001	-

β, *standardized coefficient; B, unstandardized coefficient; SE, coefficients standard error; CI, confidence interval*

posture and neuromuscular control [9]. Thus, engaging in recreational sports, may it be football or other sports, may have added advantage to balance skills even in young adults. This may be a good foundation where balance is enhanced and hopefully maintained to prevent falls and fractures in later life.

The better balance performance in footballers compared to sedentary students were apparent in the sway observed, which was measured as the CoP distance travelled. Regardless if it was during the early dynamic phase (first 5-s) or static phase (after 5 s), footballers had less sway and were able to keep their movement to a minimal throughout the UPST on their non-dominant leg with eyes closed. The CoP distance was significantly shorter despite the range of CoP values in both the x- and y-axis being similar in footballers and sedentary students. Thompson and colleagues found out that footballers have better control in their anterior-posterior and mediolateral stability than the non-athletes [23]. This finding is in line with our outcome of this study.

Lastly, footballer have their hips to thank for keeping their UPST balance time longer compared to sedentary students. Their greater trochanter movement could significantly predict up to 77% of the time to complete the UPST while using their non-dominant leg and without visual input. This means participation in recreational football may indirectly provide players with better balance and postural control from the hip that may be missing in more sedentary people. As the hip is a known controller of postural stability [24], having specific and targeted exercises to improve hip control alone for balance may be difficult. Yet, participating in football games up to three times a week could train our balance intrinsically while having fun. Krustrup and colleagues mentioned that the training component of football were able to improve the postural balance of the players [25].

5. Conclusion

This is a cross-sectional study and our results are inferential at best. Our study sample size could have been improved as well to obtain a better normal distribution of data. However, non-parametric statistical analyses are equally powerful to parametric tests when used appropriately [26].

To improve on our balance protocol, we suggest future studies consider incorporating a warm-up component before performing UPST as it may improve overall balance [27]. With regards to the maximum time of 45 s, this may be suitable for young adults but researchers may consider different times for different type of population or even consider not having a time limit to include the possibility of muscle fatigue in different types of population.

The strength of our study highlights the need to understand muscle control towards balance in early dynamic phases that will contribute to the completion of the balance task at hand. Subsequently, use of advanced technology such as a force platform to assess CoP are ideal and provides a better understanding of the intricacies that are working to provide posture and stability in individuals of all ages.

Further studies are needed to support our current findings and intervention studies in sedentary young adults will be beneficial to inform in this area. More work can be done to understand the contribution of lower limb muscles and activation especially in the more vulnerable groups like older adults and people with osteoporosis to prevent falls and fractures.

Conflict of Interest

There is no conflict of interest reported between the authors.

Acknowledgment

This project was partly funded by USM Short Term Grant (304/PPSK/61312135).

References

[1] J. Winearls, "Posture: its function in efficient use of the human organism as a total concept" Agressologie, 13, 99-101, 1972.

[2] R. Lepres, A. X. Bigard, J. P. Diard, J. F. Gouteyron, C. Y. Guezennec, "Posture control after prolonged exercise" Eur J Appl Physiol Occup Physiol.. 76(1), 55-61, 1997. https://doi.org/10.1007/s004210050212

[3] P. G. Gerbino, E. D. Griffin, D. Zurakowski, "Comparison of standing balance between female collegiate dancers and soccer players" Gait Posture, 26(4), 501-507, 2007. doi:10.1016/j.gaitpost.2006.11.205

[4] D. A. Winter, A. E. Patla, J. S. Frank, "Assessment of balance control in humans". Med Prog Technol, 16(1-2), 31-51, 1990.

[5] T. Paillard, F. Noe, "Effect of expertise and visual contribution on postural control in soccer" Scand J Med Sci Sports, 16(5), 345-348, 2006. DOI: 10.1111/j.1600-0838.2005.00502.x

[6] J. Jerosch, L. Thorwesten, H. Bork, M. Bischof, "Is prophylactic bracing of the ankle cost effective?" Orthopedics, 19(5), 405-414, 1996. https://doi.org/10.3928/0147-7447-19960501-10

[7] B. L. Riemann, R. Schmitz, "The relationship between various modes of single leg postural control assessment" Int. J. Sports Phys. Ther., 7(3), 257-266, 2012.

[8] B .A. Springer, R. Marin, T. Cyhan, H. Roberts, N. W. Gill, "Normative values for the unipedal stance test with eyes open and closed" J Geriatr Phys Ther, 30(1), 8-15, 2007

[9] M. D. Jakobsen, E. Sundstrup, P. Krustrup, P. Aagaard, "The effect of recreational soccer training and running on postural balance in untrained men" Eur J Appl Physiol, 111(3), 521-530, 2011.DOI 10.1007/s00421-010-1669-2

[10] E. S. Chumanov, C. D. Remy, D. G. Thelen, D. G, "Computational techniques for using insole pressure sensors to analyse three-dimensional joint kinetics" Comput Method Biomec, 13(5), 505-514, 2010. DOI:10.1080/10255840903350148

[11] B. Pueo, J. M. Jimenez-Olmedo, "Application of motion capture technology for sport performance analysis" Retos, 32, 241-247, 2017.

[12] S. L. Colyer, M. Evans, D. P. Cosker, A. I. Salo, "A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system" Sports Med-Open, 4(1), 24, 2018. https://doi.org/10.1186/s40798-018-0139-y

[13] E. Jonsson, A. Seiger, H. Hirschfeld, "One-leg stance in healthy young and elderly adults: a measure of postural steadiness?" Clin Biomech 19, 21(3), 688-694, 2004. doi:10.1016/j.clinbiomech.2004.04.002

[14] T. Muehlbauer, C. Mettler, R. Roth, U. Granacher, "One-leg standing performance and muscle activity: are there limb differences?" J Appl Biomech, 30(3), 407-414, 2014. https://doi.org/10.1123/jab.2013-0230

[15] N. Takeshima, N. L. Rogers, M. E. Rogers, M. M. Islam, D. Koizumi, S. Lee, "Functional fitness gain varies in older adults depending on exercise mode" Med Sci Sports Exerc, 39(11), 2036-2043, 2007. DOI: 10.1249/mss.0b013e31814844b7

[16] J. A. Kim, O. B. Lim, C. H, Yi, "Difference in static and dynamic stability between flexible flatfeet and neutral feet" Gait posture, 41(2), 546-550, 2015. https://doi.org/10.1016/j.gaitpost.2014.12.012

[17] A. Karlsson, G. Frykberg, "Correlations between force plate measures for assessment of balance" Clin Biomech, 15(5), 365-369, 2000. doi:10.1016/s0268-0033(99)00096-0

[18] A. Mansfield, E. L. Inness, "Force plate assessment of quiet standing balance control: Perspectives on clinical application within stroke rehabilitation" Rehabilitation Process and Outcome, 4, 7-5, 2015. doi:10.4137/RPO.S20363.

[19] T. Paillard, F. Noe, T. Riviere, V. Marion, R. Montoya, P. Dupui, "Postural performance and strategy in the unipedal stance of soccer players at different levels of competition" J. Athl. Train., 41(2), 172-176, 2006.

[20] R. Barone, F. Macaluso, M. Traina, V. Leonardi, F. Farina, V. Felice, "Soccer players have a better standing balance in nondominant one-legged stance" Open Access J Sports Med, 2, 1-6, 2011. doi:10.2147/OAJSM.S12593

[21] C. D. Davlin, "Dynamic balance in high level athletes" Percept Mot Skills, 98(3_suppl), 1171-1176, 2004.

[22] L. Bringoux, L. Marin, V. Nougier, P. A. Barraud, C. Raphel, "Effects of gymnastics expertise on the perception of body orientation in the pitch dimension" J Vestib Res, 10(6), 251-258, 2000.

[23] L. Thompson, M. Badache, S. Cale, L. Behera, N. Zhang, "Balance performance as observed by center-of-pressure parameter characteristics in male soccer athletes and non-athletes" Sports, 5(4), 86, 2017. doi:10.3390/sports5040086

[24] L. Allet, H. Kim, J. Ashton-Miller, T. De Mott, J. K. Richardson, "Frontal plane hip and ankle sensorimotor function, not age, predicts unipedal stance time" Muscle Nerve, 45(4), 578-585, 2012. DOI 10.1002/mus.22325

[25] P. Krustrup, P. Aagaard, L. Nybo, J. Petersen, M. Mohr, J. Bangsbo, "Recreational football as a health promoting activity: a topical review" Scand J Med Sci Sports, 20(Suppl. 1), 1-13, 2010. doi: 10.1111/j.1600-0838.2010.01108.x

[26] M. Hughes, S. M. Cooper, A. Nevill, "Analysis procedures for non-parametric data from performance analysis" Int J Perf Anal Spor., 2(1), 6–20, 2002. doi:10.1080/24748668.2002.11868257

[27] A. Daneshjoo, A. H. Mokhtar, N. Rahnama, A. Yusof, "The Effects of Comprehensive Warm-Up Programs on Proprioception, Static and Dynamic Balance on Male Soccer Player". PLoS ONE, 7(12): e51568, (2012) doi:10.1371/journal.pone.0051568

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 08-12 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

ANFIS-Based Climate Controller for Computerized Greenhouse System

Hassan Oubehar[1,*], Abdelouahed Selmani[1], Abdelali Ed-Dahhak[2], Abdeslam Lachhab[2], Moulay El Hassane Archidi[1], Benachir Bouchikhi[1]

[1]*Sensors Electronic & Instrumentation Team, Physic Department, Moulay Ismaïl University, Faculty of Sciences, Meknes, Morocco*

[2]*Modeling, Systems Control and Telecommunications Team, Department of Electrical Engineering, Moulay Ismaïl University, High School of Technology, Meknes, Morocco*

ARTICLE INFO

Article history:
Received: 07 October, 2019
Accepted: 12 December, 2019
Online: 15 January, 2020

Keywords:
Greenhouse
Climate control
Temperature
Relative humidity
ANFIS controller

ABSTRACT

The greenhouse climate system is very hard to manipulate because the variables involved are closely correlated. This study aims to enhance the regulation performance of greenhouse climate system based on adaptive neural-fuzzy inference system (ANFIS). The ANFIS is a hybrid technique that incorporates the fuzzy logic theory and artificial neural network algorithms. The employed control system has the ability to stabilize the climatic variables inside greenhouse system at required levels for crops development. The datasets utilized to train ANFIS model was obtained by implementing a fuzzy logic controller under MATLAB Simulink environment. The feasibility, reliability and robustness of the adopted strategy have been experimentally studied. The experiments conducted in real-time show that the proposed technique provides good tracking, short response time and high robustness with regard to outside perturbations and non-linear behaviors of greenhouse systems.

1. Introduction

During recent decades, computerized greenhouses become an important alternative for crops production to meet the food requirement that is increasing continuously due to the dramatic demographic growth. The quantity and quality improvements of fruits and vegetables need to maintain adequate environmental conditions under greenhouse. For that, growers have to control the main greenhouse climatic parameters that are important for plants growth [1] [2]. Nevertheless, the regulation of internal greenhouse atmosphere is a complicated process. Indeed, the different involved variables interact in a strong way. Moreover, the climate under greenhouse relies not only on the effect of external environment but also on the actions that are produced by the elements of the system through the climate control hardware [3]. During last years, various papers have investigated the smart control technologies of greenhouse production climates like PI-intelligent technique [4], fuzzy-based control systems [5-7], second-order sliding mode [8], nonlinear adaptive PID control [9] and model predictive control [10]…

*Corresponding Author: Hassan Oubehar, hassan.oubehar@gmail.com

Controllers based on conventional approaches are broadly implemented and they allow successful control for linear and time-invariant systems. Nevertheless, greenhouse is considered as complex process. In fact, it presents large nonlinear features, its parameters vary in time and also it has big reliance to the ambient environmental variables [11]. In response, the techniques of artificial intelligence (AI) were introduced in the literature to solve such issues. The main theories belonging to AI are fuzzy logic systems (FLS) and artificial neural networks (ANN). FL-based controller has many benefits for instance: the ability to reason like human being and to be easily designed based on sample dataset, so the exact mathematical model is not demanded. In addition, it has the characteristic to be adapted through changing the generated rules (there is no need to redevelopment) [12]. For the second part, ANN algorithms have great fastness and enormous learning capacity by means of a sample database and to establish nonlinear relationships between different dependent and independent variables [13]. To make use of the reasoning potential of FLS and powerful learning capacities of ANN, ANFIS algorithm has emerged for addressing the challenge of controlling complex, dynamic and non-linear systems. In this study, an ANFIS-based

control system was developed in order to promote an appropriate microclimate inside greenhouse system.

2. Schematic view of the hardware system

This study is conducted based on an automated experimental greenhouse as depicted in Figure 1. The considered prototype was home developed and installed at the Laboratory of Electronic, Automatic and Biotechnology that is a part of the Faculty of Sciences of Meknes in Morocco. A set of sensors are employed to ensure the measurements of various environmental variables involved in greenhouse system. LM35DZ sensors are utilized for measuring the outdoor and indoor temperature. The measurements of inside and outside relative humidity are performed using HIH-4000-001 sensors. To achieve control purposes, the experimental setup was equipped with an electrical heater and a variable-speed ventilator. All components are connected to an acquisition and control software units by means of a personal computer and data acquisition card NI 6024E [14-16].

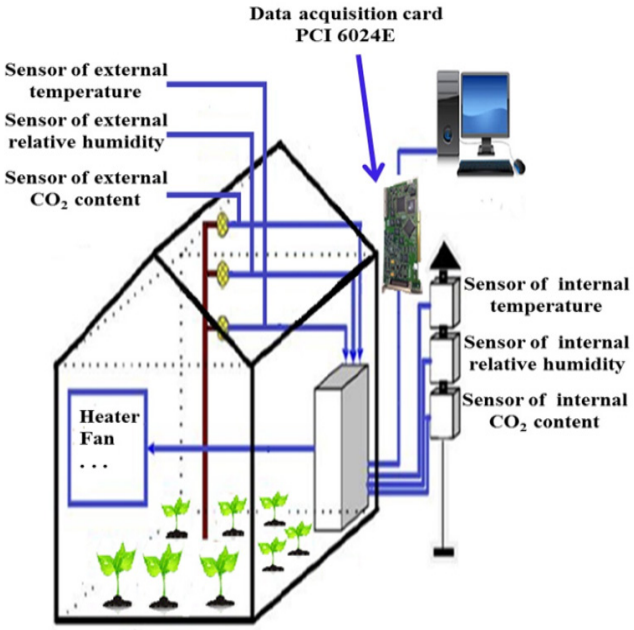


Figure 1: Experimental greenhouse system

3. ANFIS-based Control Design

ANFIS is an intelligent technique largely used by research communities for system modeling and control [17-19]. In fact, ANFIS is an efficient integration of fuzzy logic concepts and ANN algorithms. ANFIS approach are useful to build smart control technics that holds the ability of FLS to process the knowledge provided by human expert so as to imitate his actions by use of ANN learning capability [20]. ANFIS implements Sugeno-type fyzzy rules that are automatically tuned based on ANN algorithms. The fundamental principle of neuro-fuzzy systems is to create fuzzy-based model employing ANN methods where the fuzzy model parameters stand for the weights of the neural network structure [21].

Figure 2 shows the standard ANFIS architecture, where a circle denotes a fixed node, while a square represents an adaptive node. It's composed essentially of five layers (fuzzification, production, normalization, defuzzification and output layer). For simplicity, two inputs and one output are considered.

In the input layer, real values are transferred into linguistic variables. Each node is associated with a node function. The membership degree of linguistic variables is generated according to the following equation:

$$O_i^1 = \mu_{Ai}(x) \quad \text{for} \quad i = 1,2 \tag{1}$$

Here, x; Ai and μ point out input variable, fuzzy set and membership function of the i-th node respectively. If μ is generalized bell type, it is determined by three parameters {a,b,c}:

$$\text{bell}(x;a,b,c) = \frac{1}{1+\left|\frac{x-c}{a}\right|^{2b}} \tag{2}$$

The second layer is called as product layer. Each node's output is obtained by multiplying incoming signals to it with each other using the following formula:

$$O_i^2 = w_i = \mu_{Ai}(x)\mu_{Bi}(y) \quad \text{for} \quad i = 1,2 \tag{3}$$

In the third layer, each node is a fixed one known as N. The i-th node computes the normalized firing strength or normalized weight by dividing the i-th firing strength with the sum of all rules firing strengths:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad \text{for} \quad i = 1,2 \tag{4}$$

In the fourth layer all nodes are adaptive ones where the node function is given by:

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i) \quad \text{for} \quad i = 1,2 \tag{5}$$

Where $\bar{w}$ is a normalized firing strength from the previous layer and (pi, qi, ri) is the consequent parameter set of the node.

The single fixed node in the fifth layer calculates the overall output by the summation of all incoming signals using the following expression:

$$O_i^5 = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad \text{for} \quad i = 1,2 \tag{)}$$

The procedure of ANFIS design is demonstrated in Figure 3, which comprises essentially five steps. At first, the predefined dataset has been loaded into ANFIS-Editor. The following stage consists of converting the learning data into fuzzy data through a suitable membership function. Afterwards, the corresponding fuzzy rules are established based on fuzzification process. During the fourth stage, parameters of the generated model are updated

using the selected learning algorithm. At last, the defuzzification mechanism transfers the fuzzy output to real values and the constructed model is saved.

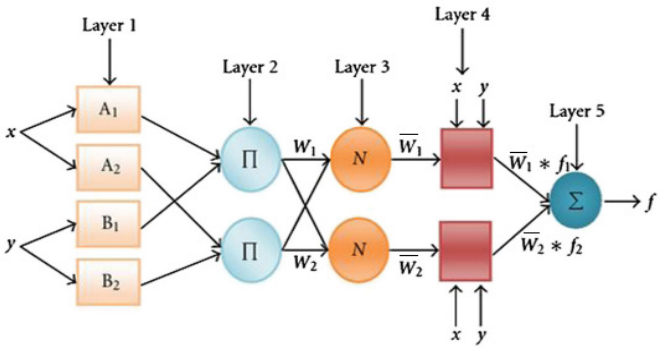


Figure 2: Basic ANFIS structure

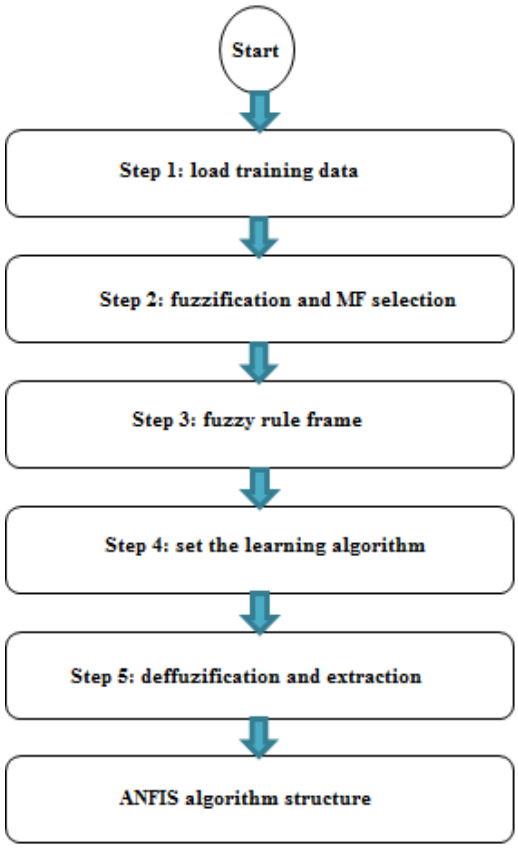


Figure 3: ANFIS flowchart

3.1. Greenhouse Climate ANFIS-based Control

The ANFIS controller developed in this works uses two inputs temperature tracking error e(T) and humidity tracking error e(RH) and a single output (voltage). A FL controller was simulated in MATLAB Simulink to provide pairs of inputs-output data. The collected database was recorded so as to be used later for training the adopted control scheme. Figure 4 illustrates the implemented control methodology. The considered control structure was applied to bring the indoor temperature (Tin) and indoor relative humidity (RHin) to desired trajectories (Ts) and (RHs). In fact, the suggested control system drives control actions via the power circuit to reach optimal climate conditions.

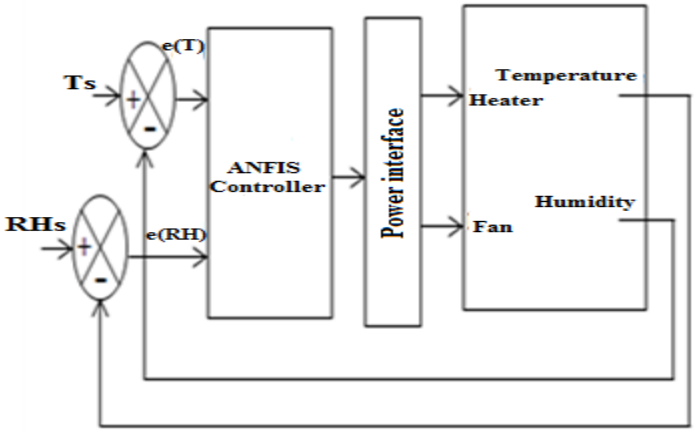


Figure 4: Schematic diagram of the proposed control process

To construct the ANFIS-based model five "generalized bell" membership functions was associated with every input as indicated in Figure 5 and Figure 6. The output variable was deffuzified using weighted average technique.

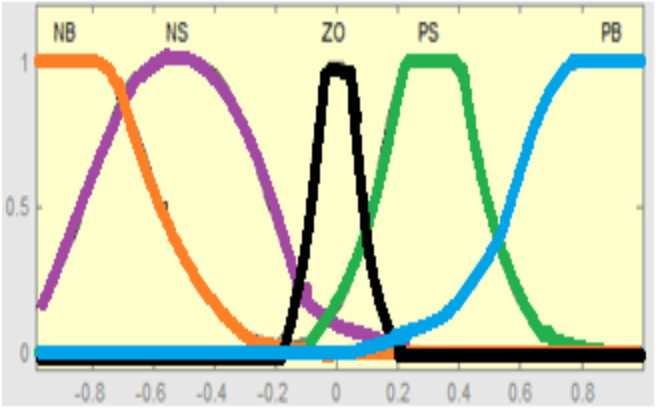


Figure 5: First input fuzzification

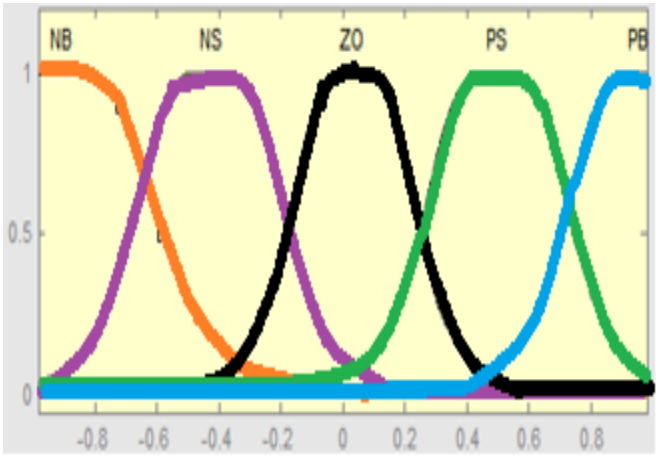


Figure 6: Second input fuzzification

The ANFIS model was created based on hybridization of two methods which consist of gradient technique to compute input parameters, and least square approach to calculate output function parameters. The architecture of the generated ANFIS controller is shown in Figure 7. The corresponding fuzzy rules are highlighted

in Figure 8. We note that the training data was normalized. Hence, the input values belong to the range [-1, 1] while the output voltage is between 0 and 1.

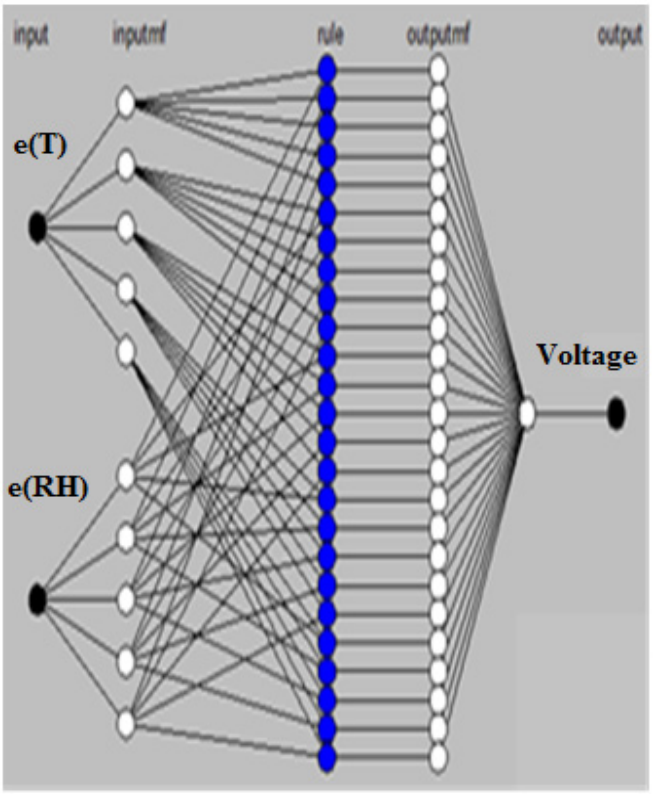


Figure 7: Structure the proposed control Model

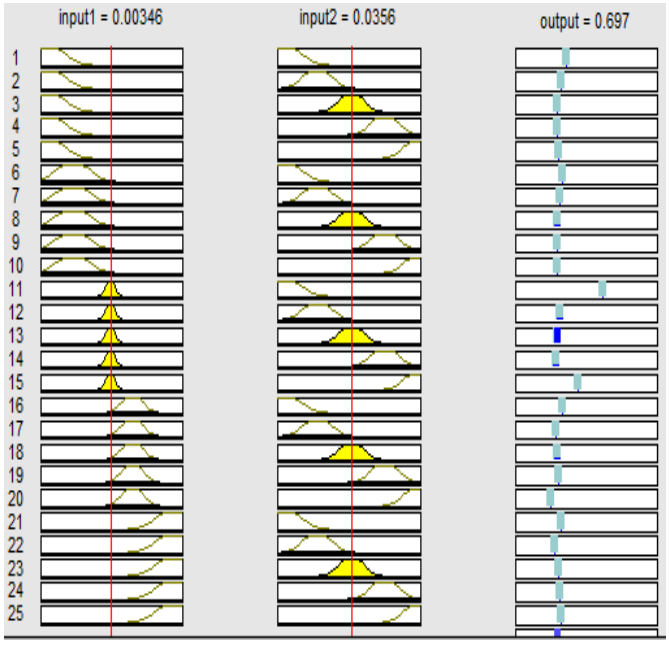


Figure 8: Fuzzy rules viewer

4. Experimental Results

The main experiments and performed measurements in real time are displayed in Figure 9 and 10. As it observed the indoor temperature (Tin) and indoor relative humidity (HRin) follow closely the predefined set points (Ts) and (RHs). The inside temperature remains close to the set point regardless of external temperature variations (Text). This proves that the developed control strategy is robust relating to external perturbations. As for the inside relative humidity, it is kept at required levels after few minutes. It is clearly obvious that the designed control system allows good regulation performance thanks to its robustness regarding disturbances due to the outside weather and system uncertainties. In fact, it permits to realize an accurate and quick response.

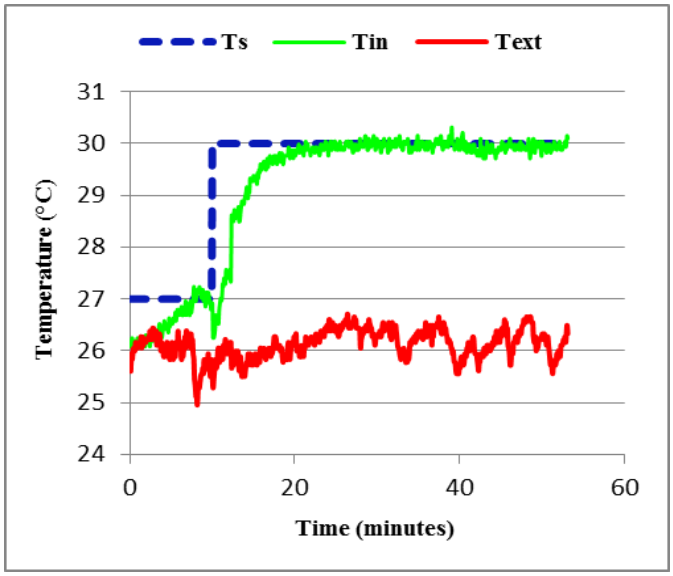
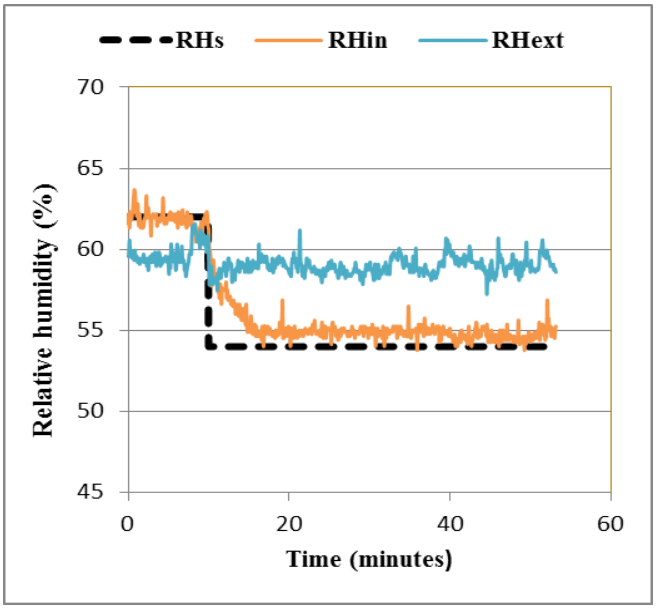


Figure 9: Evolution of internal and external temperatures



5. Conclusion

This work has developed an automation system on regulating different keys parameters of environmental greenhouse based on ANFIS algorithm. The proposed control strategy was successfully validated in real time. The experimental measurements demonstrate that the suggested technique is able to track the

reference trajectories quickly and without overshoot. It is clear that the used control method guarantees an accurate, robust and fast regulation of greenhouse indoor temperature and relative humidity to create optimal conditions for crops.

References

[1] H. Oubehar, A. Selmani, A. Ed-Dahhak, A. Lachhab, M.H. Archidi, B. Bouchikhi, "Design and real time implementation of ANFIS controller for greenhouse climate" in 2018 International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), Kenitra, Morocco, 2018.

[2] M.A. Márquez-Vera, J.C. Ramos-Fernández, L.F. Cerecero-Natale, F. Lafont, J.F. Balmat and J.I. Esparza-Villanueva, "Temperature control in a MISO greenhouse by inverting its fuzzy model," Computers and Electronics in Agriculture, vol. 124, pp. 168–174, 2016.

[3] N. Bennis, J. Duplaix, G. Enea, M. Haloua and H. Youlal, "Greenhouse climate modelling and robust control," Computers and Electronics in Agriculture, vol. 61, pp. 96-107, 2008.

[4] Y. El afou, L. Belkoura, M. Outanoute, A. Rahali, A. Lachhab, B. Bouchikhi, "Feedback Techniques Using PID and PI Intelligent For Greenhouse Temperature Control" International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering,vol. 3, pp. 9779-9792, 2014.

[5] M. Guerbaoui, A. Ed-dahhak, Y. El Afou, A. Lachhab, L. Belkoura and B.Bouchikhi, "Implementation of Direct Fuzzy Controller in Greenhouse based on Labview" International Journal of Electrical and Electronics Engineering Studies, Vol.1 No.1, pp.1-13, September 2013.

[6] M.A. Márquez-Vera, J.C. Ramos-Fernández, L.F. Cerecero-Natale, F. Lafont, J.F. Balmat, and J.I. Esparza-Villanueva, "Temperature control in a MISO greenhouse by inverting its fuzzy model," Computers and Electronics in Agriculture, vol. 124, pp. 168–174, 2016.

[7] R. Ben Ali, S. Bouadila and A. Mami, "Development of a Fuzzy Logic Controller applied to an agricultural greenhouse experimentally validated," Applied Thermal Engineering, vol. 141, pp. 798–810, 2018.

[8] H. Oubehar, A. Ed-Dahhak, A. Selmani, M. Outanoute, A. Lachhab, M. Guerbaoui, M. H. Archidi and B. Bouchikhi, " High-Order Sliding Mode Control of Greenhouse Temperature" Indonesian Journal of Electrical Engineering and Computer Science, vol. 4, no. 3, pp. 548 -554, December 2016.

[9] S. Zeng , H. Hu, L. Xu and G. Li, "Nonlinear Adaptive PID Control for Greenhouse Environment Based on RBF Network, " Sensors, vol. 12, pp. 5328-5348, 2012.

[10] L. Chen, S. Du, Y. He, M. Liang, D. Xu, "Robust model predictive control for greenhouse temperature based on particle swarm optimization," Info Proc. Agri., in press.

[11] H. Hu, L. Xu, B. Zhu, and R. Wei, "A compatible control algorithm for greenhouse environment control based on MOCC strategy" Sensors, vol. 11, no. 3, pp. 3281-3302, 2011.

[12] L. Dambrosio, "Data-based fuzzy logic control technique applied to a wind system" Enregy Procedia, vol. 126, pp. 690-697, 2017.

[13] H.K. Ghritlahre, R.K. Prasad, "Application of ANN technique to predict the performance of solar collector systems," Renewable and Sustainable Energy Reviews, vol. 84, pp. 75–88, 2018.

[14] A. Ed-Ddahhak, A. Lachhab, L. Ezzine, and B. Bouchikhi, "Performance evaluation of a developing greenhouse climate control with a computer system," AMSE Journal Modelling, Measurement and Control, vol. 68, no. 1, pp. 53-64, 2007.

[15] A Ed-Dahhak, M. Guerbaoui, Y. El Afou, M. Outanoute, A. Lachhab, L. Belkoura, and B. Bouchikhi, "Implementation of fuzzy controller to reduce water irrigation in greenhouse using LabVIEW," International Journal of Engineering and Advanced Technology Studies, vol. 1, no. 2, pp. 12-22, 2013.

[16] M. Guerbaoui,Y. El Afou, A. Ed-Dahhak, Lachhab A, and B. Bouchikhi, "PC-Based automated drip irrigation system. International Journal of Engineering Science and Technology, " vol. 5, no. 1, pp. 221-225, 2013.

[17] C.E. Lachouri, M.M. Lafifi, K. Mansouri and A. Belmeguenai, "Adaptive Neuro-Fuzzy Inference Systems for Modeling Greenhouse Climate" International Journal of Advanced Computer Science and Applications, vol. 7, no. 1, pp.96-100, 2016.

[18] N. Mathur, I. Glesk and A. Buis, " Comparison of adaptive neuro-fuzzy inference system (ANFIS) and Gaussian processes for machine learning (GPML) algorithms for the prediction of skin temperature in lower limb prostheses, " Medical Engineering and Physics,vol. 38 pp. 1083–1089, 2016.

[19] D.M. Atia and H.T. El-madany, "Analysis and design of greenhouse temperature control using adaptive neuro-fuzzy inference system, Journal of Electrical Systems and Information Technology," vol. 4, pp. 34–48, 2017.

[20] M. Zare and M. Koch, "Groundwater level fluctuations simulation and prediction by ANFIS- and hybrid Wavelet-ANFIS/Fuzzy C-Means (FCM) clustering models: Application to the Miandarband plain," Journal of Hydro-environment Research , vol. 18, pp. 63–76, 2018.

[21] B. Haznedar and A. Kalinli, "Training ANFIS structure using simulated annealing algorithm for dynamic systems identification," Neurocomputing, vol. 302, pp. 66–74, 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 6, 10-18 (2019)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

A Proposal of TCP Fairness Control Method for Two-Host Concurrent Communications in Elastic WLAN System Using Raspberry Pi Access-Point

Rahardhita Widyatra Sudibyo[1], Nobuo Funabiki [*,1], Minoru Kuribayashi[1], Kwenga Ismael Munene[1], Md. Manowarul Islam[1], Wen-Chung Kao[2]

[1]*Department of Electrical and Communication Engineering, Okayama University, Okayama, Japan.*

[2]*Department of Electrical Engineering, National Taiwan Normal University, Taipei, Taiwan.*

ARTICLE INFO

Article history:
Received: 26 July, 2019
Accepted: 30 September, 2019
Online: 20 November, 2019

Keywords:
WLAN
TCP throughput fairness
Raspberry Pi
PI controller
Test-bed
Dynamic hosts

ABSTRACT

The IEEE802.11n based Wireless Local Area Networks (WLANs) have been extensively deployed due to the flexible coverage, the easy installation, and the lower cost. To reduce the energy consumption while increasing the performance, the elastic WLAN system has been studied, such that it can dynamically change the network configuration according to traffic demands. As well, the test-bed has been implemented and for the access point (AP), Raspberry Pi is used as a portable, energy conservation, and powerful computing device. Our test-bed measurements with a single AP of two concurrently communicating hosts found the unfairness throughput results, due to the TCP windows size became different among them. To overcome this drawback, in this paper, we propose the TCP fairness control method for two concurrently communicating hosts in the elastic WLAN system. By controlling the delay at the packet transmission, the slower host will obtain more transmission opportunities than a faster host. The delay is firstly calculated by the received signal strength (RSS) from every host. After that, the delay is controlled by the PI controller to balance the both throughputs. For evaluations, we execute the proposal in the elastic WLAN system test-bed and carry out extensive measurements, where the TCP throughput fairness is achieved.

1 Introduction

The IEEE802.11n based *Wireless Local Area Networks (WLANs)* have been widely adopted around the world due to the characteristics of the flexible coverage, the simple installation, and the low cost [2]. WLAN provides the Internet access by wireless medium and offers a lot of benefits such as mobility, reliability, and portability. Hence the popularity of WLAN is increasing in government offices, private companies, organizations for accessing the internet.

In WLAN, hosts are mostly non-uniformly located [3], and communicating hosts or traffics tends to fluctuate unpredictably [4, 5] according on the period time and in a week. Furthermore, conditions of network devices and communication links may be influenced by certain factors, such as device failures, power shortages, weather changes, or bandwidth controlled by their authorities [6].

Under such circumstances, we have examined the *elastic WLAN system* that dynamically optimizes the network configuration based on demands in a network, to reduce the energy consumption during the time to improve the performance [7, 8]. In addition, we have developed the *elastic WLAN system test-bed* using *Raspberry Pi AP*. *Raspberry Pi* is a card-size, single-board computer that can solve a variety of practical problems requiring computation or networking abilities [9], additionally, it is equipped with the built-in wireless network interface (NIC) supporting IEEE802.11n.

In WLAN, the *fairness* of the throughput quality among the hosts is necessary to ensure the fair *quality of services (QoS)* for the users [10]. Hence, fairness problems in WLAN have been explored extensively [11]-[14]. Since a large number of network applications in the Internet adopt the *transmission control protocol (TCP)*, the TCP fairness is exceedingly critical.

Nevertheless, our preliminary measurements using the elastic WLAN system test-bed have revealed that the TCP throughput is not superior among the two concurrently communicating hosts associated with the same AP when they are located at different positions from the AP. It is assumed that this unfairness was caused by differences in the *TCP window size* and the *modulation and coding*

[*]Corresponding Author: Nobuo Funabiki, Dep. of Electrical and Communication Engineering, Okayama University, Okayama, Japan, Email: funabiki@okayama-u.ac.jp
This paper is an extension of work originally presented in 2nd International Conference on Communication Engineering and Technology (ICCET 2019) [1].

scheme (MCS) among them. Therefore, the packet transmission interval of the farthest host hereafter becomes longer, and the interval for the nearest host becomes shorter.

In this paper, we propose the *TCP fairness control method* for two concurrently communicating hosts in the elastic WLAN system. The *transmission delay* is implemented at the AP in the packet transmission to the nearer host. After that, it can be predicted that the nearer host will decrease the throughput, and the farther host will enhance it by obtaining more transmission opportunities. This delay is firstly calculated by the *received signal strength (RSS)* from hosts. After that, it is dynamically controlled using the *PI controller* [15] to achieve the desired throughput fairness.

For performance evaluations, the proposal is implemented in *Raspberry Pi* AP and conduct experiments to verify the effectiveness using the test-bed with static and dynamic hosts in different network fields.

The remainder of this paper is structured as follows: Section 2 shows the preliminaries of the works. Section 3 demonstrates the TCP fairness control method. Section 4 and 5 evaluates the proposal through test-bed experiments. Section 6 discusses the observation result. Finally, Section 7 concludes this paper with future works.

2 Preliminaries

In this section, we introduce our preliminary work to relate this paper.

2.1 Elastic WLAN System

The *elastic WLAN system* has been designed and developed to control and manage the number of active APs in the network field according to the traffic demand. Figure 1 shows the test-bed topology of the system.

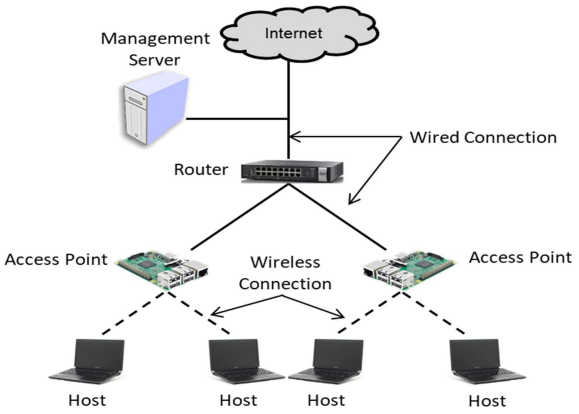


Figure 1: Elastic WLAN system topology.

The server has the administrative access to all the network devices including the APs, so that it can control them by executing the *active AP configuration algorithm* by the following procedure:

- The server examines the network devices and collects the necessary information for the active AP configuration algorithm.
- Then, it executes the active AP configuration algorithm and the output of the algorithm contains the minimum number of APs, the AP-host associations, and the allocated channels.
- Finally, it implements the output of the previous step by deactivating or activating the selected APs, changing the requisite host associations, and allocating the channels.

2.2 Software AP Configuration

In this paper, we adopt *Raspberry Pi 3*, that uses *Raspbian OS*, *Broadcom BCM2837 SOC*, *LPDDR2-900MHz 1GB SDRAM*, *10/100Mbps Ethernet*, *IEEE 802.11b/g/n wireless NIC*, *Bluetooth 4.1 classics/low energy*, and *400MHz video core IV GPU* [9]. The procedure to configure the *Raspberry Pi 3* device to function as a Wi-Fi AP is as follows:

- First, *hostapd* is installed by the command as follows:

```
#apt-get install hostapd
```

- Then, the */etc/hostapd/hostapdraspi.conf* configuration file is modified with the SSID and PASSWORD. An example of the configuration file is provided as follows:

```
interface=wlan0
ssid=RaspberryAP
wpa_passphrase=raspberry
channel=13
```

- Next, the absolute path is set to run *hostapd* while booting the system using the following command:

```
DAEMON_CONF="/etc/hostapd/
   hostapdraspi.conf"
```

- The static IP address in the em wlan0 interface is assigned by modifying file */etc/network/interfaces* as follows:

```
auto wlan0
iface wlan0 inet static
address 172.24.2.10
netmask 255.255.255.0
network 172.24.2.0
```

- Finally, install the DHCP server for assigning the dynamic IP addresses to the hosts.

2.3 *Jain's Fairness Index*

There are various definitions have been proposed for the fairness index. Among them, *Jain's fairness index* [16] is the representative one for the TCP fairness:

$$F(n) = \frac{\left(\sum_{i=1}^{k} x_i\right)^2}{k \times \sum_{i=1}^{k} x_i^2} \quad (1)$$

Here k represents the total number of connections and x_i does the throughput for the i-th connection. When all the connections receive

the same throughput, this fairness index becomes 1, which means that all the connections are 100% fair. As the disparity increases, the fairness will decrease.

2.4 Unfairness Problem in Concurrent Communication

We initially measured the TCP fairness among concurrently communicating multiple links with a single AP using the elastic WLAN system test-bed with the various distances between host. We conducted the experiments in *Asahi Riverbed* as an outdoor environment. By using *Homedale* [17], it is confirmed that all interfering signals have been removed from this environment.



Figure 2: Outdoor measurement field.

Figure 2 indicates the locations of the AP and two hosts in the measurements. The position of the host *H2* is changed from 0*m* to 30*m* with the 5*m* step while the position of *H1* is fixed at the 0*m* distance from the AP. Then, the throughput and RSS are measured at both *H1* and *H2*. For each measurement, *iperf software* [18] is used to generate the traffic from both hosts at the same time to the AP.

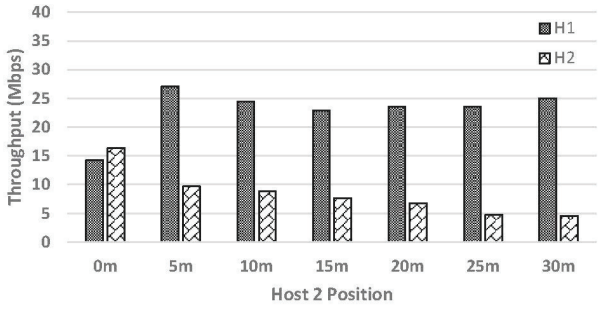


Figure 3: Throughput at different locations for *H2*.

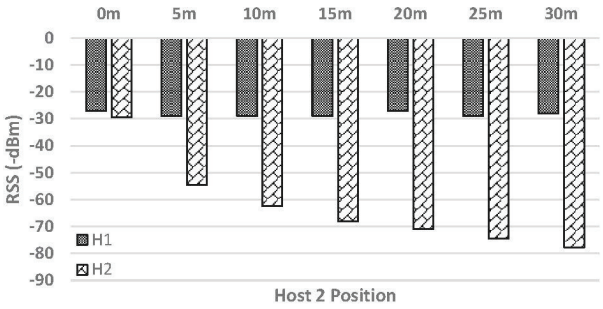


Figure 4: RSS at different locations for *H2*.

Figure 3 indicates throughput measurement results, which are similar at both hosts when the *H2* distance from the AP is 0*m*. However, as the *H2* distance increases, the throughput difference will increase accordingly. The throughput of *H2* decrease due to the smaller RSS (receiving signal strength). Nevertheless, the throughput of *H1* will increase, although RSS at *H1* is constant as shown in Figure 4. This is considered as an inspiring result, that is to say, it seems that the nearer host to the AP will take higher transmission opportunities than the further host, which leads to the unfairness of the TCP throughput performance among them.

3 Proposal of TCP Fairness Control Method

In the section, we describe the proposal of the TCP fairness control method in the concurrent TCP communications of two hosts with a single AP.

3.1 Overview

Our test-bed experiments results unfold that the two-host concurrently communicating with a single AP causes the unfairness in the throughput performance, due to the different packet transmission intervals among them. Figure 5a demonstrates the packet transmission intervals between the near and far hosts. The throughput unfairness happens by the differences in the MCS and TCP window size. For the near host, the TCP windows size will be larger because of faster MCS compare to the far host. Then, it will occupy the far larger bandwidth during the communications compare to the other host, thus the packet transmission interval between them increases.

To overcome the issue, the *transmission delay* is intentionally introduced in the packet transmission of the AP to the near host. By minimizing the transmission bandwidth of the link with the near host, it will give more transmission bandwidth to the link to the far host. To achieve the throughput fairness, this delay should be controlled such that the intervals of packet transmission become equal, as shown in Figure 5b.

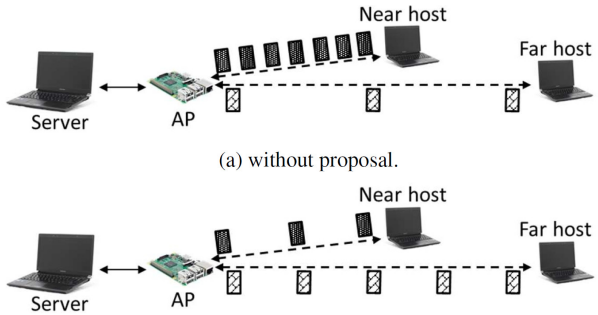


Figure 5: Background of TCP fairness control method.

The transmission delay is initially calculated by collecting the RSS at the AP from the near host and from the far host, because the measured RSS influences the MCS. After that, the transmission delay is controlled dynamically by the PI control during communications to attain the fairness accurately in the network.

3.2 Initial Delay Calculation

The initial transmission delay $D(0)$ is computed from the difference of the measured RSS between the two hosts:

$$D(0) = \frac{RSS_{near}}{-a}\left(\frac{RSS_{near}}{RSS_{min}}\right)^2 e^{b(RSS_{near}-RSS_{far})} \quad (2)$$

The parameter values are obtained from our experimental results in [19]. Where, $RSS_{min} = -88dBm$ from [20] is considered as the minimum RSS that a host can receive packets successfully from a *Raspberry Pi* AP, RSS_{near} does the measured RSS at the AP from the near host, RSS_{far} does the measured RSS from the far host, and the constant parameter a and b are needed to be set depending on the environment. Here, $a = 8$, $b = 0.15$ for the outdoor field, and $a = 9$, $b = 0.17$ for the indoor field are used, respectively. If the RSS is smaller than RSS_{min}, the host cannot receive any packet successfully.

3.3 Dynamic Delay Optimization by PI Controller

In this stage, to achieve the target fairness index, the delay is dynamically controlled by the *PI controller* [15]. The delay is optimized dynamically as follows:

$$D(n) = K_P \times (F_{tar} - F(n)) + K_I \times \sum_{i=0}^{n} (F_{tar} - F(i)) \quad (3)$$

Here, $D(n)$ represents the delay, $F(n)$ does the measured fairness index at the n-th time-step, and K_P and K_I does the gain for the P control and for the I control, respectively. Where, $F_{tar} = 0.97$, $K_P = 3$, and $K_I = 0.8$ are used. It is noted that the PI controller periodically changes the delay with the $20sec$ time-step. In the system implementation, the equation in (3) is modified as follows:

$$D(n) = D(n-1) + K_P \times (F(n-1) - F(n)) + K_I \times (F_{tar} - F(n)) \quad (4)$$

The measured fairness index $F(n)$ is calculated by:

$$F(n) = \frac{\left(x_{far}(n) + x_{near}(n)\right)^2}{2 \times \left(x_{far}(n)^2 + x_{near}(n)^2\right)} \quad (5)$$

where $x_{far}(n)$ and $x_{near}(n)$ represent the measured throughput of the far host and that of the near host at the n-th time-step, respectively.

3.4 Test-bed Implementation

Here, we show the implementation of our proposal in the elastic WLAN system test-bed.

3.4.1 RSS Measurement

We collect the measured RSS of every host by using the command as follows at the AP:

```
$ sudo iw dev wlan0 station dump |
    egrep "Station|signal:"
```

where the AP does not require to access to the host. Since the RSS continually fluctuates, RSS is measured for 30 seconds, and their average is used in Eq. (2).

3.4.2 Transmission Delay Application

Next, the transmission delay is applied to the packet transmission from the AP to the nearest host by the following procedure:

- Remove any delay previously assigned at the interface *wlan0*:

```
$ sudo tc qdisc del root dev wlan0
```

- Assign each delay at the packet transmission from the AP to every host, namely $d1$ and $d2$:

```
$ sudo tc qdisc add dev wlan0 root
    handle 1: prio bands 3
$ sudo tc qdisc replace dev wlan0
    parent 1:1 netem delay $d1
$ sudo tc qdisc replace dev wlan0
    parent 1:2 netem delay $d2
```

- Enforce every delay to the packet transmission to the nearest host and the far host that is specified by the IP address:

```
$ sudo tc filter add dev wlan0
    protocol ip parent 1: u32 match
     ip dst $IPaddressofnearhost
    flowid 1:1
$ sudo tc filter add dev wlan0
    protocol ip parent 1: u32 match
     ip dst $IPaddressoffarhost
    flowid 1:2
```

Here, the IP address of each host is automatically set by the following program:

Linux commands for near host detection

```
#/bin/bash
# for check total connected hosts
01: sudo hostapd_cli all_sta |
grep "dot11RSNAStatsSTAAddress="
while [ "$i" -le "$totalConnectedHost" ]
do
# for collecting the IP address
02: arp -a | grep $mac
# merge MACAddress, IP Address and RSS
03: echo $mac" " $ip" " $rss >> data.tmp
done
# sort descending data.tmp by rss value
04: sudo sort -nrk 3 data.tmp> sort.tmp
# 1st line in sort.tmp is the nearest host
```

3.4.3 Dynamic Delay Optimization

Firstly, the initial delay is obtained to the nearest host from the server by the following step:

- Explore the IP and MAC addresses of the connected hosts.
- Request the AP to measure the RSS of the hosts for 30sec and send back to the server.
- Find the average of the measured RSS.
- Compute $D(0)$ by Eq. (2).
- Assign $D(0)$ as the transmission delay to the nearest host.

Then, the delay is optimized dynamically by the following steps at every 20sec:

- Measure the throughput of the hosts for 20sec using *iperf*.
- Compute $F(n)$ by Eq. (5).
- Update $D(n)$ by Eq. (4).
- Set $D(n)$ as the transmission delay to the faster host at the AP.

4 Evaluations with Static Hosts

In this section, we evaluate the proposal through experiments when all the hosts are stationary.

4.1 Experiment Setup

In our experiments, we generate TCP traffic from two concurrently communicating hosts to the server to measure the throughput. Figure 6 illustrates the network setup of the server, the AP, and the hosts. The server is connected with a 100$Mbps$ wired connection to the AP, and the host is connected to the AP through the IEEE802.11n 2.4Ghz wireless link. The traffic is generated from the host to the server using *iperf*. The buffer size and the TCP window size are set to 8KB and 477KB, respectively.

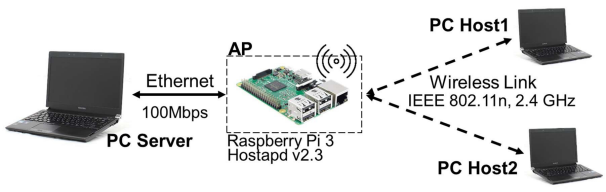


Figure 6: TCP link configuration.

Table 1 reveals software and hardware specifications for our experiments. Three laptop PCs are used for the two client hosts and the one server, and one *Raspberry Pi* is configured for the AP operating on channel 13.

Table 1: Hardware and software specifications.

server and hosts	
type	R731/B, Toshiba Dynabook
operating system	Linux (Ubuntu 14)
RAM	4GB DDR3-1333MHz
processor	Intel Core i5-2520M, 2.5Ghz
software	Iperf 2.0.5
access point	
type	Raspberry Pi 3
operating system	Linux (Raspbian)
RAM	LPDDR2 900MHz 1GB
processor	BCM2837 1.2Ghz, Broadcom ,
NIC	BCM43438, Broadcom
software	hostapd

4.2 Experiments in Outdoor Field

First, the experimental results are provided in the outdoor field.



Figure 7: Outdoor field.

4.2.1 Experiment Field

As the outdoor field, experiments are conducted in *Asahi Riverbed* shown in Figure 7. There is no interfering signal, which is confirmed by using *Homedale*. The height of the AP and the hosts are 135cm and 70cm, respectively. During experiments, the distance of $H1$ from the AP is set 0m, 5m, 10m, and 15m. Then, for each $H1$ distance, $H2$ is located at 0m, 5m, 10m, 15m, 20m, 25m, and 30m from the AP. Finally, the throughput experiments are carried out in three methods: *1) without proposal*, *2) initial delay only*, and *3) proposal*.

4.2.2 Throughput Results

Figure 8 illustrates the throughput results of the two hosts for *1) without proposal*, where the unfairness appears between $H1$ and $H2$. Figure 9 shows the results for *2) initial delay only*, where the fairness of each host is improved. Figure 10 shows the results for *3) proposal*, where the fairness is achieved at any position by implementing our proposal. Figure 11 shows the overall throughput comparison at each $H1$ location. Nevertheless, with the proposal, the overall throughput has been reduced by 2.78% on average.

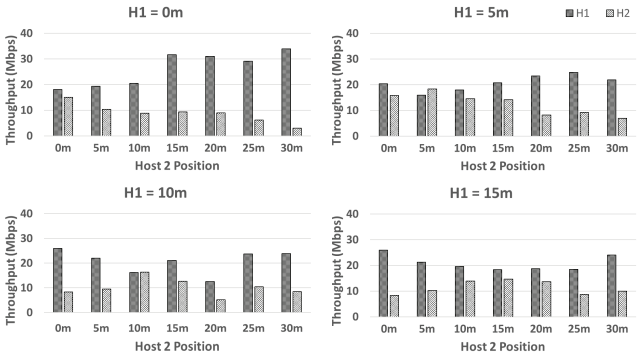


Figure 8: Results of throughput in outdoor field for *1) without proposal*.

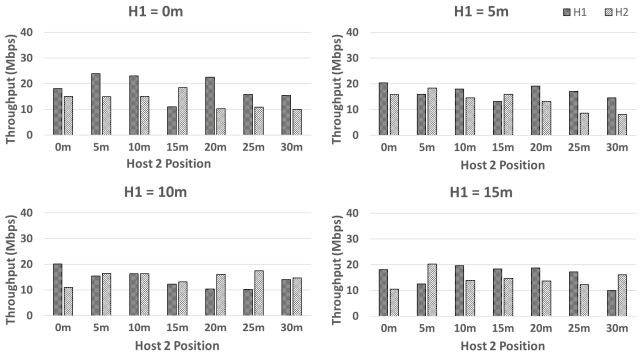


Figure 9: Results of throughput in outdoor field for *2) initial delay only*.

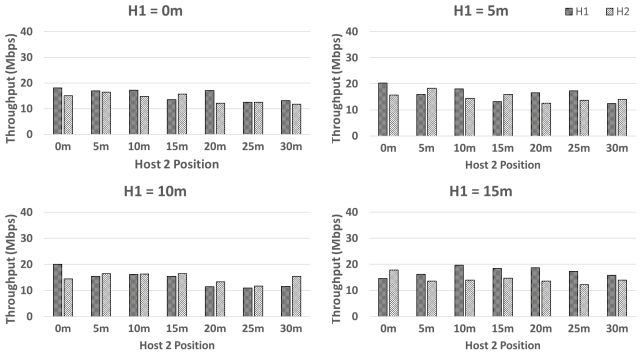


Figure 10: Results of throughput in outdoor field for *3) proposal*.

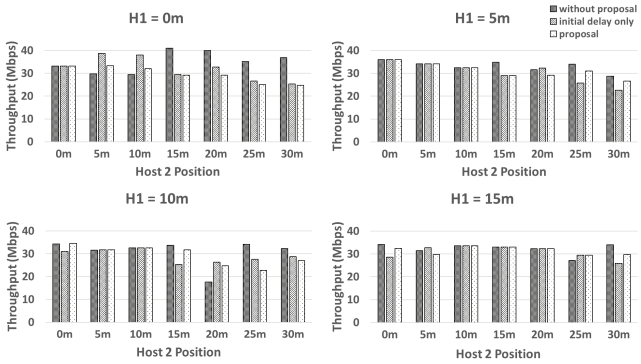


Figure 11: Results of overall throughput in outdoor field.

4.2.3 Fairness Index Results

Figure 12 demonstrates the *Jain's fairness index* result. From the Eq. (1) known 1 is the finest fairness index. For *1) without proposal*, any result appears to be lower than 1. Nevertheless, for *3) proposal*, the fairness index will generally become close to 1. It is found that the proposal presents better achievement in the fairness index than the comparisons.

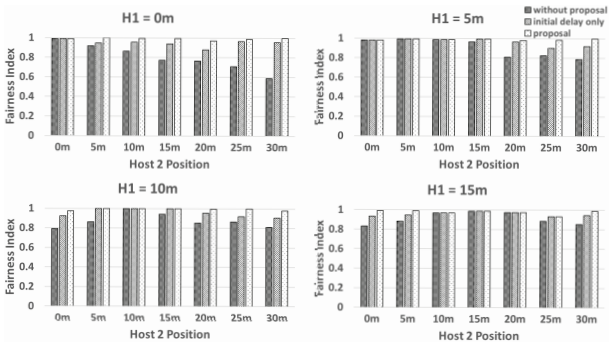


Figure 12: Observed TCP fairness index in outdoor field.

4.3 Experiments in Indoor Field

Second, experimental results in the indoor field are provided.

4.3.1 Experiments in Two-room Case

The two rooms in the 3rd floor of Engineering Building #2 at Okayama University are used in the experiment. Figure 13 illustrates the layout of this field.

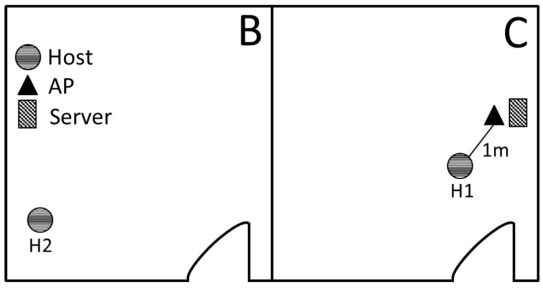


Figure 13: Experimet field for two-room case.

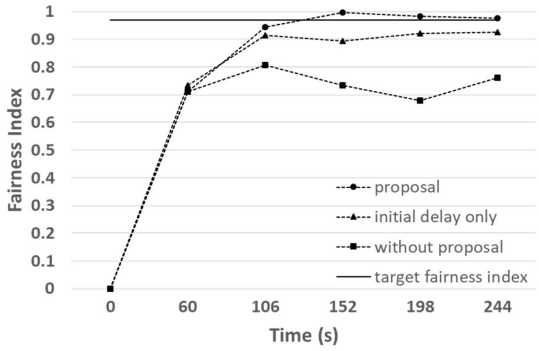


Figure 14: TCP fairness index results in indoor field for two-room case.

Figure 14 shows the fairness index results. The system collected the RSS and other necessary data during the first 60*sec* to measure the throughput. Next, it was observed that the fairness index for each method is about 0.7. Then after 130*sec*, fairness is achieved for *3) proposal*. Conversely, *1) without proposal* and *2) initial delay only* could not achieve it.

4.3.2 Experiments in Four-room Case

Then, the four rooms on the same floor are used in the experiment. Figure 15 illustrates the layout of this building that has a complex structure of multiple rooms and walls. This figure shows the position of the AP (triangle), the server (square), and the possible position of the hosts (circle). Here, the AP and the host are configured similarly to the outdoor field. $H1$ is located at A1 and B2 in the indoor field. Then, for each $H1$ position, $H2$ is moved from A1 to D4, respectively.

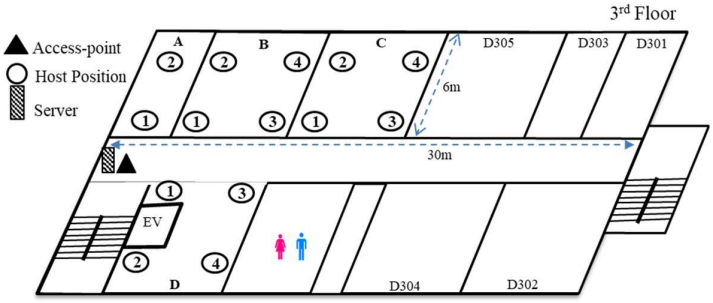


Figure 15: Indoor field.

Figure 16 illustrates the throughput results of the two hosts for *1) without proposal*. The unfairness appears between $H1$ and $H2$. Figure 17 shows the results for *2) initial delay only*, where the fairness is slightly improved. Figure 18 shows the results for *3) proposal*, where the fairness is achieved by the proposal. Figure 19 shows the overall throughput comparison at each $H1$ location. However, by applying the proposal, the overall throughput has diminished by 14.83% on average.

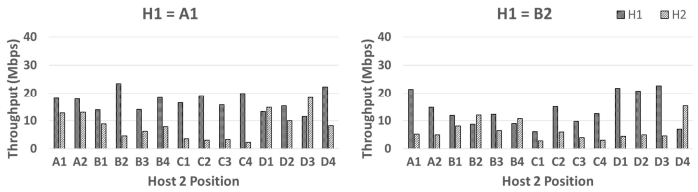


Figure 16: Results of throughput in indoor field for *1) without proposal*.

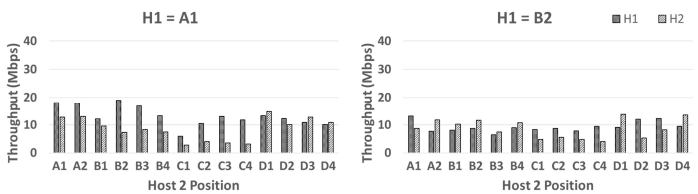


Figure 17: Results of throughput in indoor field for *2) initial delay only*.

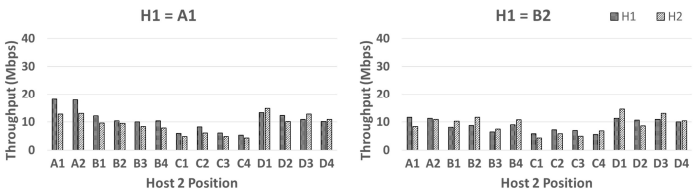


Figure 18: Results of throughput in indoor field for *3) proposal*.

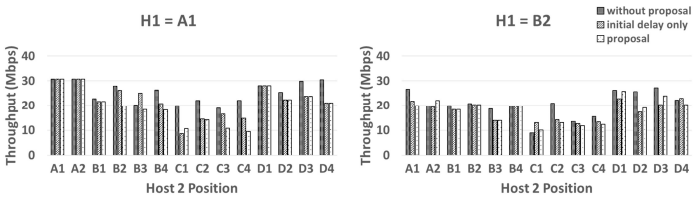


Figure 19: Results of overall throughput in indoor field.

Figure 20 shows the fairness index result of the four-room case. For *1) without proposal*, the fairness index appears to be smaller than 1. Nevertheless, for *2) initial delay only*, it increases significantly. Furthermore, for *3) proposal*, it is going to be nearly 1. Thus, the effectiveness of the dynamic delay optimization using the PI controller is confirmed.

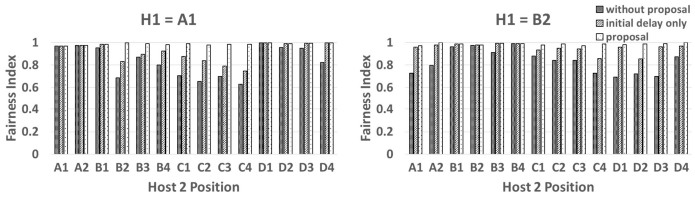


Figure 20: TCP fairness index results in indoor field for four-room case.

5 Evaluations with Dynamic Hosts

Next, we evaluate the TCP fairness control method through experiments when one host is dynamically connected and disconnected to the AP.

5.1 Experiment Setup

The topology in Figure 13 and the devices/software in Table 1 are used in this experiment. Figure 21 indicates the 20*min* scenario of dynamically connecting and disconnecting one host. Initially, $H1$ is connected to the AP. At 0*min*, *iperf* TCP traffics are generated for 5*min*. At 5*min*, $H2$ is connected to the AP, and the proposed method is started to run manually. Then, the server will collect the RSS and measure the throughput at $H1$ and $H2$. Afterward, at 10*min*, $H2$ is disconnected from the AP, and the server measures the throughput of $H1$. Lastly, at 15*min*, both $H1$ and $H2$ are connected.

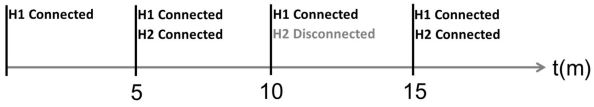
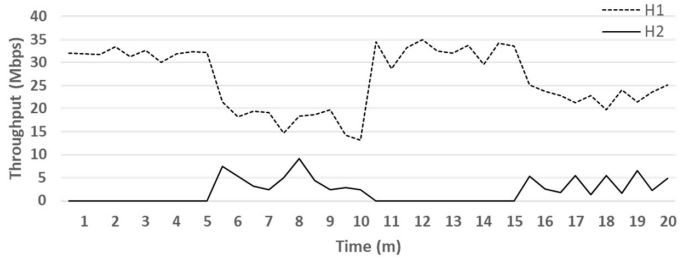
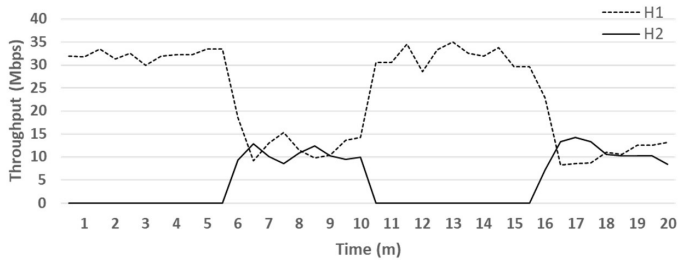


Figure 21: Time scenario for dynamically connecting/disconnecting host.



(a) Results of throughput without proposal.



(b) Results of throughput with proposal.

Figure 22: Individual throughput of $H1$ and $H2$ with dynamic hosts.

5.2 Results

Figure 22 shows the individual throughput for $H1$ and $H2$. At $0min$, where only $H1$ is connected to the AP, the individual throughput is about 33Mbps. Next, at $5min$, where both hosts are connected, the individual throughput of $H1$ decreases due to the concurrent communication. Here, *with proposal* can achieve the throughput fairness between two hosts, where it is around $12Mbps$ for each host. On the other hand, *without proposal* cannot achieve it. Thus, our proposal has proved to reach the throughput fairness between two concurrently communicating hosts with a single AP, even if a host is continually connected and disconnected to the AP.

6 Discussions

In this section, we discuss the summaries of experiment results in the previous sections and future works.

The TCP fairness control method is proposed in this paper so that any host communicating with the same AP can enjoy the similar throughput regardless of the relative position from the AP. This service fairness is important in designing and managing a local-area network, including the wired and wireless LANs. To verify the effectiveness of the proposal, evaluations with static and dynamic hosts are conducted using real devices.

First, the experiment results with static hosts in Section 4 show the improvement of the fairness index from 0.85 to 0.98 on average by the proposal, whereas the average overall throughput is dropped by 2.78% in outdoor fields and by 14.83% in indoor fields on average. Since the possible throughput of WLAN is increasing rapidly due to the advancements of wireless communication technologies, these drops of the overall throughput are acceptable.

Second, the experiment results with dynamic hosts in Section 5 show that the fairness index is improved from 0.69 to 0.97 on average by the proposal, whereas the average overall throughput is dropped by 3.21% on average. This small drop of the overall throughput is acceptable.

From these results, it is confirmed that the TCP fairness control method is practical and useful in achieving the throughput fairness among the two concurrently communicating hosts with the AP.

The weakness of the proposal is the use of the throughput measurement software (*iperf* in this paper). Every host in WLAN must install the client software, which can be an obstacle for practical use. In future works, we will investigate a method of estimating the throughput without installing this software at a host.

Besides, the proposed method is based on empirical works with a lot of experiments using real devices. The theoretical analysis of two-host concurrent communications and the model formulation for simulations will be also in future works. This model should accurately estimate the throughput for a given transmission delay at the packet transmission from the AP to a host when two hosts are concurrently communicating with the single AP.

7 Conclusion

This paper proposed the *TCP fairness control method* for two concurrently communicating hosts in the elastic WLAN system where the delay is implemented in the packet transmission of the faster host. Initially, the delay is estimated by collecting the RSS of the hosts at the AP. Then, it is dynamically controlled using the PI controller to achieve throughput fairness between the hosts. Extensive experiments with static and dynamic hosts confirmed the effectiveness of the proposal. In future works, the proposal will be extended to three or more concurrently communicating hosts and carry out experiments in different topologies and network fields.

Acknowledgments

This work is partially supported by JSPS KAKENHI (16K00127).

References

[1] R. W. Sudibyo, N. Funabiki, M. Kuribayashi, K. I. Munene, M. M. Islam, and W.-C. Kao, "A TCP fairness control method for two-host concurrent communications in elastic WLAN system using Raspberry Pi access-point," in Proc. 2nd Int. Conf. Commun. Eng. Tech. (ICCET 2019), pp. 76-80, 2019.

[2] B. P. Crow, I. Widjaja, J. G. Kim, and P. T. Sakai, "IEEE 802.11 wireless local area networks," IEEE Commun. Mag., vol. 35, no. 9, pp. 116-126, Sept. 1997.

[3] K. Mittal, E. M. Belding, and S. Suri, "A game-theoretic analysis of wireless access point selection by mobile users," Comput. Commun., vol. 31, no. 10, pp. 2049-2062, Jan. 2008.

[4] M. Balazinska and P. Castro, "Characterizing mobility and network usage in a corporate wireless local-area network," in Proc. Int. Conf. Mob. Syst. Appl. Ser., pp. 303-316, 2003.

[5] D. Kotz and K. Essien, "Analysis of a campus-wide wireless network," Wirel. Netw., vol. 11, no. 1-2, pp. 115-133, Jan. 2005.

[6] F. Nadeem, E. Leitgeb, M. S. Awan, and S. Chessa, "Comparing the life time of terrestrial wireless sensor networks by employing hybrid FSO/RF and only RF access networks," in Proc. Int. Conf. Wirel. Mob. Commun., pp. 134-139, 2009.

[7] M. S. A. Mamun, M. E. Islam, N. Funabiki, M. Kuribayashi, and I-W. Lai, "An active access-point configuration algorithm for elastic wireless local-area network system using heterogeneous devices," Int. J. Network. Comput., vol. 6, no. 2, pp. 395-419, July 2016.

[8] M. S. A. Mamun, N. Funabiki, M. E. Islam, and W.-C. Kao, "A channel assignment extension of active access-point configuration algorithm for elastic WLAN system and its implementation using Raspberry Pi," Int. J. Network. Comput., vol. 7, no. 2, pp. 248-270, July 2017.

[9] Raspberry Pi, "The official website of the Raspberry Pi project," http://raspberry.org. Accessed 19 Feb., 2019.

[10] S. Pilosof, R. Ramjee, D. Raz, Y. Shavitt, and P. Sinha, "Understanding TCP fairness over wireless LAN," in Proc. IEEE INFOCOM, April 2003.

[11] T. Nandagopal, T. Kim, X. Gao, and V. Bharghavan, "Achieving MAC layer fairness in wireless packet networks," in Proc. ACM Mobicom, Aug. 2000.

[12] N. Blefari-melazzi, A. Detti, A. Ordine, and S. Salsano, "A mechanism to enforce TCP fairness in 802.11 wireless LANs and its performance evaluation in a real test-bed," in Proc. World. Wirel. Mob. Multi. Net., pp. 1-7, 2007.

[13] B. A. H. S. Abeysekera, T. Matsuda, and T. Takine, "Dynamic contention window control mechanism to achieve fairness between uplink and downlink flows in IEEE 802.11 wireless LANs," IEEE Trans. Wirel. Commun., vol. 7, no. 9, pp. 3517-3525, 2008.

[14] E. Park, D. Kim, H. Kim, and C. Choi, "A cross-layer approach for per-station fairness in TCP over WLANs," IEEE Trans. Mob. Comput. vol. 7, no. 7, pp. 898-911, 2008.

[15] K. Astrom and T. Hagglund, "PID controller: theory, design, and tuning, 2nd Ed.", Inst. Socie. America, 1995.

[16] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system," East. Res. Lab. (DEC), vol. 38, 1984.

[17] Software Verzeichnis development, Homedale WLAN Monitor, http://www.the-sz.com/products/homedale. Accessed 19 Feb., 2019.

[18] Iperf, The ultimate speed test tool for TCP, UDP and SCTP, https://iperf.fr/. Accessed 27 Feb., 2019.

[19] R. W. Sudibyo, N. Funabiki, M. Kuribayashi, K. I. Munene, M. M. Islam, and W.-C. Kao, "A TCP fairness control method for concurrent communications in elastic WLAN system using Raspberry Pi access-point," IEICE Tech. Rep., NS2018-175, pp. 96-100, Dec. 2018.

[20] Cypress, "CYW43438, Single-Chip IEEE 802.11 b/g/n MAC/Baseband/Radio with Integrated Bluetooth 4.2," 2016.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 16-27 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Novel Hybrid Method for Segmentation and Analysis of Brain MRI for Tumor Diagnosis

Kapil Kumar Gupta[*,1, 2], Namrata Dhanda[1], Upendra Kumar[3]

[1]*Department of Computer Science and Engineering, Amity University, Lucknow, India*

[2]*Department of Computer Science and Engineering, Shri Ramswaroop Memorial University, Barabanki, India*

[3]*Department of Computer Science and Engineering, Institute of Engineering and Technology, Lucknow, India*

ARTICLE INFO

Article history:
Received: 01 January, 2020
Accepted: 11 April, 2020
Online: 03 May, 2020

Keywords:
Brain Tumor
Image Segmentation
Fuzzy C-Mean Clustering
Adaptive mean and variance based filter
Noise Removal

ABSTRACT

It is difficult to accurately segment brain MRI in the complex structures of brain tumors, blurred borders, and external variables such as noise. Much research in developing as well as developed countries show that the number of individuals suffering tumor of the brain has died as a result of the inaccurate diagnosis. The proposed article, a novel hybrid method improves segmentation accuracy. The proposed research includes three basic steps. In the first step, the adaptive filter based on mean and local variance is utilized for noise removal in the input images. It helps in de-noising to a different orientation and scale, creates numerous responses for all components in the medical images while preserving the edges. In the second step, the development of a hybrid method takes place. It is the combination of extended K-mean clustering and fuzzy C-mean clustering. The purpose of the research is to develop a hybrid segmentation structure of single-channel T1 MR Images for multiform benign and malignant tumors. It removes the limitation of prefixed cluster size which helps in improving the segmentation accuracy by reducing the sensitivity of the clustering parameters. In the third step, the morphological non-linear operation performed for the removal of the non-tumor part. The proposed approach is evaluated against various statistical parameters such as mean, standard deviation, entropy, correlation, homogeneity, smoothness and variance. The parameters result predicts a greater balance between the automated tumor areas extracted by radiologists with the tumor areas extracted by the proposed method. The findings show that the proposed hybrid method achieves a 98% level of segmentation accuracy.

1. Introduction

Advanced 3-D medical imaging techniques produce highly efficient and reliable images [1]. Segmentation of the medical image is the method of automatically or semi-automatically detecting boundaries within a 2D or 3D image. A significant challenge in segmenting medical images is the increased variation of medical images. First and foremost, there are significant modes of variation in human anatomy. Furthermore, a number of different modalities like X-ray, CT, MRI, Microscopy, PET, SPECT, Endoscopy and OCT are used in the processing of medical images [2]. Normal cells develop regulated in the human brain, whereas new cells replace damaged cells, tumor cells produce uncontrolled in the brain, which is not yet understood. Benign or malignant can be primary brain tumors. Brain tumors are heterogeneous and extremely variable in size, place, shape, and appearance. Automatic benign and malignant brain tumor segmentation is a difficult task. Some MR imaging artifacts also increase the complexity level in the tumor segmentation [3]. The slow growth of a benign brain tumor is certain and rarely diffuses, while its cells are not malignant. Figure 1 display below containing different types of benign MRI tumor images.

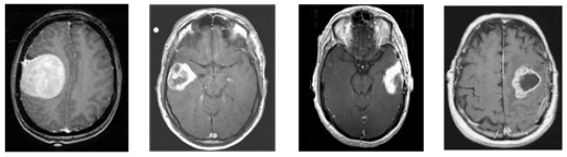


Figure 1: Brain MRI images containing different types of benign brain tumor

[*]Kapil Kumar Gupta, Amity University, Shri Ramswaroop Memorial University, kapilkumargupta2007@gmail.com

Figure 2 display containing different types of malignant MRI tumor images.

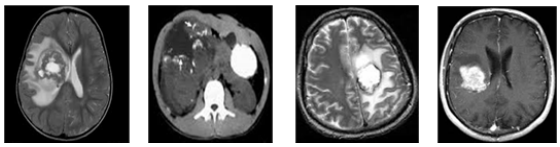


Figure 2: Brain MRI images containing different types of malignant brain tumor

The tumor, causing normal tissue damage, compression and elevated intra-crane pressures can have an effect on the brain [4]. Symptoms like headache, dizziness and visibility problems can vary by type, size, and location of the tumor in the brain.

In this paper, Image segmentation relates to pixel classification or grouping, so that each class or group represents an object that is viewed differently. Different characteristics are used for this purpose to discriminate between one item and another [5]. Texture, boundaries, edges, and color are some of the characteristics most commonly used to differentiate between different objects [6, 7]. Edge is a significant image feature, so edge detection plays an important role in the segmentation. A challenging issue for edge detection of digital medical images is noise. De-noising images is a classic yet active subject as it is an essential step in many applications. Since the pixel characteristics are discontinuous, there will be comparatively evident variations between the pixel characteristics on both sides of the edges. Therefore, the fundamental concept of the edge-based segmentation algorithm is to use some techniques to discover the limits. On one side of the corners the pixels are divided into one sub-image and on the other side, the pixels are considered to belong to another sub-image. It is sensitive to noise and generally gets incomplete data [8]. Linear filters that transform the image into a constant matrix, in the presence of additive noise, will generate the image that is blurred with the poor position of features and incomplete noise reduction [9]. A lot of research has taken into consideration and conclude that the adaptive mean and variance based filter is better than other filters for de-noising medical images.

Clustering techniques are used to distinguish groups of similar objects in a multivariate data set gathered from areas such as Biomedical and Geospatial etc.. The Fuzzy C- Mean (FCM) clustering technique applies to a broad range of data in the assessment of biomedical information. For any set of numerical information, this program produces fuzzy partitions and prototypes. These partitions are helpful to support recognized substructures or to suggest substructure in unexplored information. A generalized minimum-square objective function is the clustering criterion used to aggregate subsets [10]. Extended K-means algorithms are used to separate the area of concern from the context [11].

For the assessment of Brain Tumor, the statistical parameters are implemented. Under the average data technique, statistical and mathematical parameters such as Entropy, Standard Deviation, Mean and Variance are introduced [12]. For the number of iterations, the statistical range of each parameter is calculated. The individual statistical parameter assessment is performed with its impact on MR images of the brain tumors. Factors such as intensity, artifact presence and similarities in gray-level images are also taken into account in the segmentation process.

2. Background Study

Segmentation of the magnetic resonance images relates to the pixel clusters. Each cluster represents an object that is viewed differently. Distinct characteristics are used to discriminate between one object to another. Medical image processing is always a focal point for researchers. The challenge is to predict the accurate areas of brain tumors within the given time frame. Many studies suggested in the field of brain tumor identification have been released over the past 30 years. Research has grown exponentially over the past decade. An article was written by A. Kharrat et. al. on the brain tumor recognition using the Wavelet Transform segmentation approach to break up images of MRI. In this proposed method k-means are implemented to remove the locations or tumors that are suspect. The wavelet transformation functions $\Psi_{x,y}(s)$called the wavelet family in equation 1. Equation 2 evaluates many wavelets, which are derived from the scaling of a solitary function $\Psi(s)$ called the mother wavelet [13].

$$B_{x,y}= \int_{-\infty}^{\infty} a(s)\, \Psi_{x,y}(s)ds \quad (1)$$

$$\boldsymbol{\Psi_{x,y}(s)} = \frac{\mathbf{1}}{\sqrt{\mathbf{x}}}\boldsymbol{\Psi}\left(\frac{\mathbf{s-y}}{\mathbf{x}}\right) \text{ with } x \neq 0 \quad (2)$$

The method displays images at different levels of resolution. The wavelet test also allows the images to be compact or to de-noise without any noticeable degradation. In this study, researchers achieved 21.6272 PSNR value and 17.7768 MSAD value. R. Ratan and. al. suggested a detection method for brain tumors based on multi-parameter analyzes of MR images. Segmentation of the Watershed was implemented in this process. By deciphering a force picture's inclination guide as a tallness appreciation, we get lines that seem to be edged by all accounts. The journey from the dividing lines to the corresponding catchment bowl would be discovered if the forms were a precipitation dropper. These lines of isolation are called watersheds. The shift in the watershed can be created on a grayscale by flooding method. It was used to detect brain tumors in 2D and 3D medical images [14]. A properly organized quick diagnostic of brain tumor technique was suggested by Qurat-Ul-Ain et al. In the first stage extraction of texture, features takes place while the second stage consists of the classification of brain pictures on the basis of texture feature using the hybrid base classifier. The extracted tumor region is categorized as malignant by using two-phase segmentation processes. The segmentation process in this technique involves stages of removal of the skull and extraction of tumors. The researchers achieved 99% classification accuracy [15]. M. U. Akram et al. suggested an automatic diagnostic system technique for MR images. In this approach, segmentation and fragmentation of the images are obtained in three phases. In the first phase, preprocessing is performed to evacuate the blurring and to sharpen the picture. In the second phase, thresholding based segmentation takes place and in the third phase, fragmented images are processed with morphology, and tumor cover is carried out to remove the falsified pixels. The authors achieved an average of 97% segmentation accuracy in the experiment [16]. S. Baurer et al. placed forward a distinctive approach in which tumor patients can select a healthy brain atlas for medical images. The simulation of tumor growth along with trained algorithms is used to equate a safe image of the brain with that of the patient's brain. Finally, the non-stringent identification between the mutated atlas and the patient

image is used to allow a better contrast. It provides flexibility in the segmentation and enhanced tumor simulation and tumor enhancement treatment of atlas-based brain tumor images. In this approach, iterations are repeated until 90% of the tumor volume achieved [17]. A multiform brain tumor segmentation proposed by J. Huo et. al. used a hybrid approach. In this approach, they adopted three different methods: the first one is fuzzy connectedness, the second one is grow cut, and last is a classification of voxel using the vector supporting machine. As a hybrid rule, a confidence map averaging (CMA) technique has been used. In order to capture the global "joining together" of the voxels, a fuzzy correctness frame (FC) assigns the target object flipped affinities during classification. A paired t-test was found to be greater than 5 in the results of the experiments. A. Tom et. al. Suggested a technique of segmentation using invariant geometric transform analysis to detect brain tumors. The authors proposed a unique algorithm which is the combination of translational, rotational and scaling. The feature vector is calculated using the combination of shape, position, and texture. After that, the calculation of the Euclidean distance between the input vector and the preciously stored vector takes place. Images with minimum Euclidean distance are considered whose result is above a certain level i.e. threshold level. Experimental results achieved 90% of segmentation accuracy [18]. R. Vijayarajan et.al. explain about the medical image segmentation using fuzzy C-mean clustering technique where averaging is done by fusion of principal components. Image fusion is a technique of transmitting all appropriate and complementary image information into a single composite image from multiple MRI from the same source or from different sources. In this method, source images are segmented by the FCM clustering algorithm into K-number of clusters and the segmented regions are sorted specifically for fusion regions. If the principal components for the k-clusters is $p1^i$ and $p2^i$ where, i is from 1 to K then the average of all the parts is assessed by equation 3 and 4 [19].

$$p1= (1/K) \sum_{i=0}^{k} p1^i \quad (3)$$

$$p2= (1/K) \sum_{i=0}^{k} p2^i \quad (4)$$

K is subjective to various image inputs. And p1 and p2 are fusion weights. The fused image is obtained by the following equation 5 :

$$I=p1*i1+p2*i2 \quad (5)$$

Where, i1 and i2 are two images which are going to be fused. The average quality index is around 85% which better than the other algorithm as given in the experiments.

The distinctive local, independent transformation based segmentation and classification of CT and MRI images have been submitted by M. Huang et al. In their study, the LIC method is employed to identify the segments into separate classes. Locality plays an important part in LITC's autonomous urban development, which solves problems. The locality is used to bind the anchor code to solve linear transformation issues instead of other coding techniques. On the basis of the actual patient results, the average dice similitude of the proposed whole tumor, tumor center, and tumor section are 0.84, 0.685 and 0.585, respectively [20]. The method was suggested by N. Dhanachandra et. al. for the solution of the problem of segmentation using the K-mean algorithm of clustering and subtractive clustering algorithm. In this approach, the enhancement in the image quality, partial stretching improvement is applied to the picture. The clustering subtractive approach is an information clustering technique. Therefore, the subtractive cluster is used in k-means algorithms for the division of the images. The median filter is then placed in the segmented image to eliminate any undesired part from the image. Experimental results showed that RMSE is 0.0017 and PSNR is 35.77 which is a good result for segmentation. The ACM segmentation and ANN-LM classification methods used in brain tumor MRI analysis were used by A. Shenbagarajan and others. The proposed image analysis of MRIs using the ACM method regionally used to segment and the ANN algorithm based on the LM to efficiently identify MRIs as normal brain and timorous brain. The proposed MRI image-based brain tumor evaluation should effectively address the segmentation process and the classification method for brain tumor evaluation using feature removal techniques so that the technique would lead to a better result of brain tumor diagnostics before use in medical areas [21]. M. Sornam and others focus on establishing an independent diagnostic tumor program using pictures of MR weighted T1 and T2. The beginning step is to separate the brain MR images into benign and malignant by filtering images and the k-mean algorithm. The textural and form-based extraction process is performed with Wavelet and Zernike methods in the intermediate step. The final step is to determine the difference entre benign and malignant tumors using the ELM (Extreme Learning Machine) algorithm. Researchers found that the proposed algorithm achieved 77% of segmentation accuracy [22]. Figure 3 depicts an example of wavelet and Zernike based segmentation results.

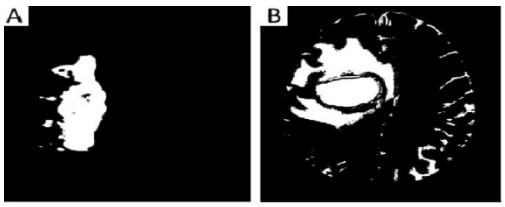


Figure 3: A – Benign, B – Malignant are the images after segmentation using k-mean [22].

In order to classify images of the MR of brain tumor, R.Ahmed et al. suggested the use of a support vector machine and ANN. For the segmentation of MRI images, the combination of tempered K-means and modified Fuzzy C-means (TKFCM) is used. TKFCM assimilates with few improvements K-means and Fuzzy C-means. A combination of Temper with K-means has identified in the brain MR images by gray level intensity. Equation 6 shall assess the gray based strength and picture complexity.

$$A(a_i, b_j) = \sum_{i=m+1}^{P+m} \sum_{j=m+1}^{Q+m} R(a_i, b_j)\, S_{PQ(resize)} \quad (6)$$

Equation 7 gives the Temper window.

$$S_{PQ} = \sum_{i=m+1}^{P-m} \sum_{j=m+1}^{Q-m} R(a_i, b_j) \quad (7)$$

Tempe-based image matrix is given with G gray-level intensity number and B bin numbers used for image R (a_i, b_j) temper detection. Where, m is defined as m= (winsize -1)/2. Application

of exact temper values, rank and column is obtained by placing the desired temper.

$$L_g = \sum_{i=1}^{D} \sum_{j=1}^{G} A(a_i , b_j) || a_i - d_j ||^2 \quad (8)$$

$$L_p = \sum_{j=1}^{Q} \sum_{i=1}^{G} (V_{ij}) * (h_{ij})^2 \quad (9)$$

The description with which the tumor area could be located as follows (8) and (9) equation:

$$L_{gp} = \oint_{d}^{\cdot} (L_g , L_p) \quad (10)$$

P and Q are row and column of A (a_i, b_j) and d is contour value. Cluster number, cluster center, and data point's number are identified by G, Q, and D respectively. The proposed study evaluates the sample data with 97.37% classification accuracy [23]. X. Ma et.al. proposes a noise removal method in SAR images. In this method, shift-invariant K- means singular value decomposition (SVD) with guided filter is used. The entire technique is made up of two steps. The first stage is to deal with the noisy image, the K-SVD shift, and the original de-noisy images. In the second step, the guided filtration is performed for the initial de-noisy image. Researchers achieved an average of 31.4040 PSNR over 6 images in the given method [24]. The automated brain tumor identification technique suggested by A. Mukaram et. al. is composed of the picture and separating two principal sections. Smoothing of the image can be accomplished through the use of a median filter, followed by an image enhancement method that can be accomplished through Sobel edge detection. For the Segmentation purpose, the K-means pillar algorithm is used. Pillar K-mean algorithm involves pillar pixel selection for efficient segmentation [25]. Hanafy M. Ali presented a study on MRI images for noise removal. MRI images are usually susceptible to noise such as the Gaussian noise, salt, pepper, and spatula noise. Therefore, having a precise brain picture is an extremely severe activity. A precise brain image is very important for further diagnosis. To remove noise, a median filter algorithm is used. Input images are combined with Gaussian noise and salt and pepper noise at various levels. To solve the problem of noise reduction it will be introduced a combination of the proposed Middle filter (MF), adaptive Middle Filter (AMF) and adaptable Wiener Filter (AWF). The filters are used to eliminate the additive noise from images of the MRI [26]. S. Riaz et. al. applied a unique technique on images having large scale data. This incorporates rough Fuzzy C-mean clustering with the methodology of the neural network. The main concept is to first create an original cluster core, then update image clusters and representations jointly during preparation, using a single layer of clustering by multi-layers of CNN. Fuzzy rough C-mean (FRCM) will be used for upgrading the forward pass cluster centers, while Stochastic Gradient Decent (SGD) will be used to upgrade CNN parameters. This rough collection of low and limited approaches addresses uncertainty, vacuity, and incompleteness in the definition of clusters and fuzzy sets which make it possible to effectively manage overlapping partitions in a noisy environment. The experimental result predicts that the given method has an average of 91% clustering accuracy among 4 different datasets [27]. Exhaustive research carried by A. Zotin et. al. used in brain tumor pictures to detect edges. In this technique, you look for a set of the fuzzy cluster by searching iteratively. A related fuzzy membership matrix specifying the membership level of the kth element is prepared in the first set of vectors of the ith cluster. Equation 11 calculates the objective function.

$$f(D, C) = \sum_{i=1}^{n} \sum_{k=1}^{m} d \cdot (x_m - c_n)^2 \quad (11)$$

Where, d is the membership value, x_m is the mth cluster and c_n is the centre of the nth cluster. After dividing the image into a series of homogenous clusters with the fuzzy C-mean (FCM) algorithm the canny edge detector is implemented. The authors achieved 98.4% segmentation accuracy in 2 images [28, 29]. Proposed A Robust segmentation on MRI and field correction, incorporated into a clustering model through local contextual information by Z. Zhang and J. Song. In this study, the precise segments of brain MRI, which is corrupted by noise and intensity homogeneity, are used to produce a new robust clustering of local context information (RC-LCI). In order to produce the respective anisotropic weight to update the present core pixel and eliminate noisy pixels, the weighting method incorporating local contextual data was used to construct a pixel in the principal pixel region. A multiplying structure consisting of a real image product and a bias field could then efficiently segment the brain RMI and analyze the bias domain. In terms of the coefficient of Jaccard similarity, the results obtained by RC_LCI were increased by 0.195 +/- 0.125 [30]. A unique algorithm proposed by F. Han for the detection of high-speed moving locations in a noisy environment. In conjunction with the Wavelet Coefficients tree and Lipschitz Expertise Noise Property, the definition of the neural network activation function is agreed on for the activation evaluation technique. Then it is possible to retain the important wavelet coefficients. The non-important coefficients were also excluded by the process of testing the independent coefficients. High-frequency data can be preserved during the wavelet transformation by raising the disappearance of wavelet filters which is useful for edge sensing [31]. A. Nyma et. al. presented a research paper on a hybrid technique for MRI segmentation. In this study, they adopt the vector median filter for the removal of noise. Thereafter, the Otsu threshold is used to define the homogenous regions of the input image in the first coarse segmentation process. An improved FCM is used to separate brain MR pictures in several segments that use the optimal deletion factor in the given data set for perfect clustering. The experimental result showed that it achieves around 96% of segmentation accuracy [32]. M.Z. Abderrezak et. al. discuss a new hybrid approach for the analysis of brain MRI. In this approach, the nonlocal median filter is used for the preprocessing. For the post-processing, authors propose a new model that is based on geometrical transformation. The segmentation efficiency is up to the mark for the tumor and the multiple detections of sclerosis [33]. R.Ahmad et. al. discussed a different approach for solving brain tumor problem. The combined SVM and ANN classification techniques were adopted by the researchers. In order to segment an image, temper based KMFCM is used. The numbers of clusters are assigned more than the standard K-mean algorithm. Automatically modified FCM membership eradicates tumor field contouring issues. The statistical characteristics obtained from segmented images are used to identify and distinguish tumors from normal SVM brain MRI images. The classification of benign tumors is based on 2 sets of features and ANN is used for four stages of malignant. In this method, 97.37% classification accuracy achieved.

3. Proposed Hybrid Method

On the basis of the historical research and result analysis of the different segmentation methods for brain tumor diagnosis, the authors proposed a unique hybrid method for automatic brain tumor detection. The framework of the proposed method is given below:

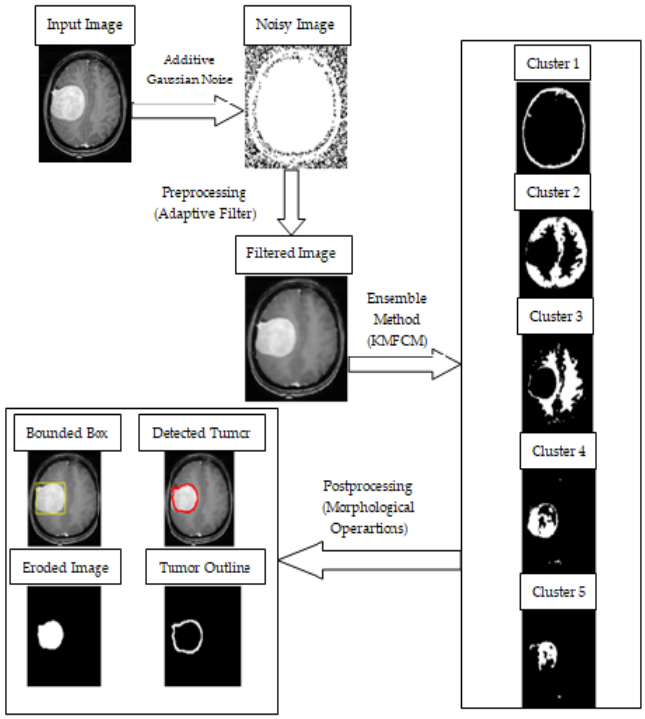


Figure 4: The framework of the proposed novel hybrid method for brain tumor segmentation

As shown in Figure 4, this article primarily divides the segmentation algorithm into three components.

Step 1: Preprocessing: Input the brain image from the Kaggle database [34]. Add Gaussian noise with different levels like 0.005, 0.007, etc. In this step noise removal is takes place with the help of adaptive mean and variance bas filter. By smoothing the MR images, it eliminates noise. This filter also reduces the variability in the intensity of MR images between one pixel to another.

Step 2: Applying Hybrid Method: In this step, a hybrid method which is the combination of modified K-mean clustering and fuzzy C- mean clustering (KMFCM) is applied. Segmented image with 5 different clusters are shown in figure 4. Brain tumor clustering and segmentation is done according to a threshold value.

Step 3: Post-processing: In the post-processing morphological operations are applied to the finding exact tumor region. For this purpose, the bounded box is used to show closed tumor areas. Also, the exact prediction of brain tumor some more eroded images and tumor outline images is added.

3.1. Preprocessing

On the degraded images containing Gaussian noise, the adaptive filter is applied. The mean and variance are the two statistical measures depending on a locally adaptive filter with a specified region of m × n window. The Function of adaptive filter is given below:

$$f(x,y) = g(x,y)- (\sigma^2_{noise}/ \sigma^2_{local})(g(x,y)-\mu) \quad (12)$$

Where, g(x, y) is the pixel value of the image at position (x, y), σ_{noise} is the variance of the overall noisy image, σ_{local} is the variance of the local region and μ is the local mean. The overall noise removal process is explained below with the help of an example:

- Define a 3 × 3 window size.
- Consider an input image B with additive Gaussian noise level=0.005. Noisy image is a 218 × 180 matrix having values from 0 to 255.
- Appending 0's to all four sides of the matrix B. The resultant matrix is C of order 222 × 184.
- Calculate the local mean (μ) and variance by sliding 3 × 3 window on the entire matrix C. Local mean is the mean of window of a matrix of 3 × 3.
- Calculate local variance (σ_{local}) with the given below equation:

$$\sigma_{local} = \text{mean}(\text{window}^2)-(\text{mean}(\text{window}))^2 \quad (13)$$

- The noise variance (σ_{noise}) of the whole image is calculated by taking the average of all local_variance.
- If ($\sigma_{noise} > \sigma_{local}$) then $\sigma_{local} = \sigma_{noise}$.
- The Filtered image is calculated by the following equation:

$$I_{filtered} = B-(\sigma_{noise}^2/ \sigma_{local}^2) \times (B- \mu) \quad (14)$$

Noisy image and the filtered image is shown in the figure below:

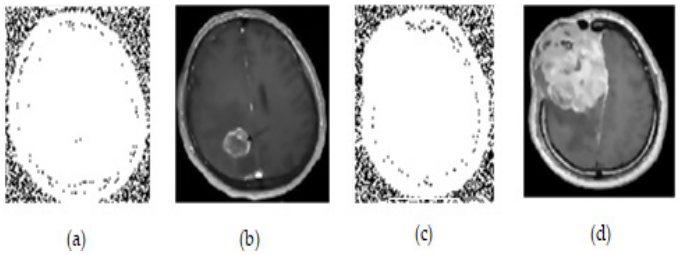


Figure 5: (a) Noisy Benign brain tumor image (b) Filtered image of an after processing adaptive filter (c) Noisy Malignant brain tumor image (d) Filtered image of c after processing adaptive filter

3.2. Applying Hybrid KMFCM Method

The proposed hybrid method is the combination of the modified K-mean clustering technique and Gaussian based fuzzy C- mean clustering technique. In the first stage, the extended K-mean is applied for deterministic cluster centroid initialization to prevent overfitting and then use the Fuzzy C-means algorithm for enhancing the classification capability. The extended K-mean algorithm is based on the standard K-mean algorithm having the capability of deterministically initializing the centroids. Maximizing the distance between the original cluster centroids is the fundamental principle of the extended K-means algorithm for initializing cluster centroids. This technology allows cluster

centroids to be initialized deterministically and resolve the shortcomings of the K-means algorithm linked to their initialization instability. Given below is the method of initialization of an extended K-mean algorithm:

- Select a sample pixel as the first initialized cluster centroid from the information set randomly.
- Select the centroids of the all k- clusters and compute the distance between sample pixel (i.e. the point within the cluster) and cluster centroid.

$$d_i = \sqrt{((x_c-x_i)^2+ (y_c-y_i)^2)} \quad (15)$$

Where, (x_c, y_c) is the centroid of the cluster and (x_i, y_i) is the ith sample pixel.

- After finding the distances between cluster points with the remaining pixel values, extract the point with minimum distance.
- Compute the probability of remaining sample pixel by using equation 16

$$P(\text{pixel})= d(x')^2/\Sigma_{x\in R}\, d(x)^2 \quad (16)$$

Where, x is the sample pixel within the set of R pixels. d(x) is the distance between the sample pixel x to closest cluster center and d(x') is the distance between the nearest pixels. By probability, select the sample with the biggest probability as the new centroid cluster.

- Repeat the above method until the determination of k cluster centroids.
- Now applying, fuzzy C-mean clustering for further updating in the cluster center.
- Let, X_i is the input cluster centers obtained from extended k-mean clustering.
- FCM partitions input information set X_i, where i=1 to R, is assigned to different clusters by assigning membership values to the set of points.
- Calculate the membership values with the help of objective function given below:

$$F_m = \sum_{i=1}^{R}\sum_{j=1}^{c} \mu_{ij}^m (x_i - c_j)^2 \quad (17)$$

Where, m>1, μ_{ij} is the degree of membership of cluster input data set x_i in the jth cluster having cluster centre c_j.

- k is the number of steps. Iteratively optimizes the objective function and updates the degree of membership μ_{ij} and the c_j cluster centers.

$$\mu_{ij} = \left(\sum_{k=1}^{c} \left(\frac{x_i-c_j}{x_i-c_k}\right)^{2/(m-1)}\right)^{-1} \quad (18)$$

$$c_j = \left(\sum_{i=1}^{R} \left(\mu_{ij}^m \times x_i\right) \Big/ \sum_{i=1}^{R} \mu_{ij}^m\right) \quad (19)$$

- Calculate the value of $\mu_{diff} = \mu_{ij}^{(k+1)} - \mu_{ij}^{k}$. If the μ_{diff} is less than ε then terminates. The value of the ε is the fixed value between 0 and 1 throughout the computation of the hybrid method.

The results after applying the hybrid method are given below in figure 6:

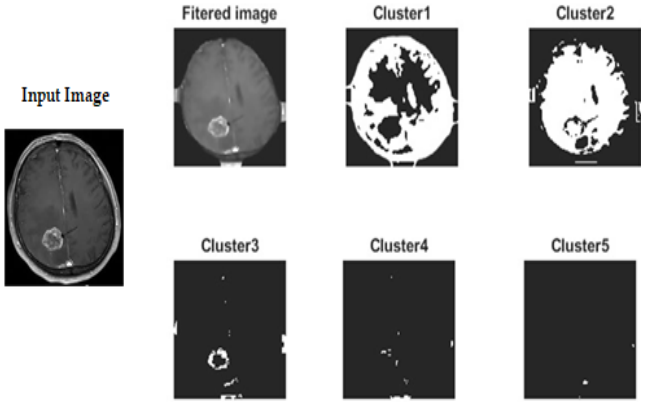


Figure 6: Clusters numbered 1 to 5 are shown after processing input image to the proposed KMFCM method

3.3. Post-processing

Morphological operations such as erosion and dilations applied to the results to improve the accuracy of segmented areas.

- Erosion is a method in which the structuring element is translated across the image domain. In a binary picture, erosion shrinks or thins items. A structuring element controls the way and magnitude of shrinking.
- The erosion by a structuring factor S of the binary image I generates a new binary image $I_{eroded} = I \ominus S$ with those positions (x, y) where the structuring element S matches the input image I originated, i.e. I(x,y) = 1 suited I and 0 otherwise, this process will be repeated for all the image pixels.
- Dilation is an operation in the binary image that develops or thickens items. A structuring component controls the way and magnitude of shrinking. Dilation is a method in which the structuring element is translated through the image domain.

$$\oplus$$

The results after applying the morphological operations are given below in figure 7:

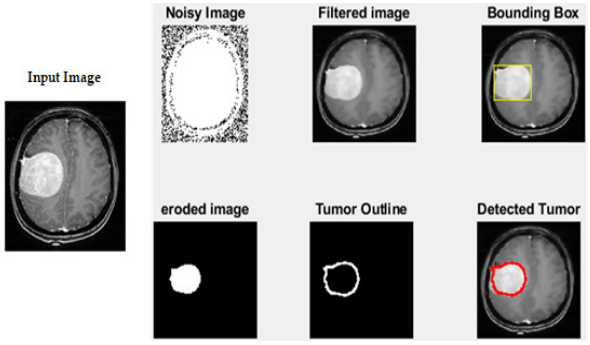


Figure 7: Eroded image, dilated image (Tumor Outline) and detected tumor are shown after processing input image to the morphological operation.

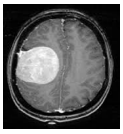
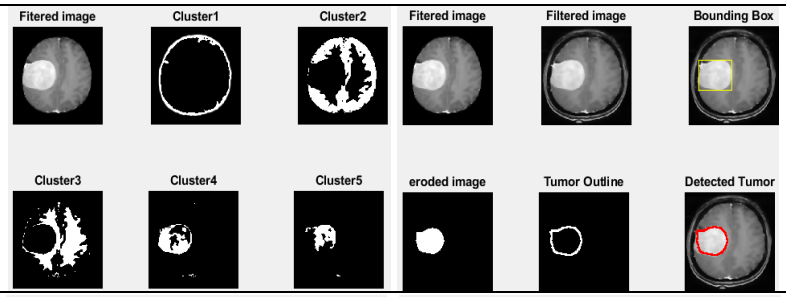
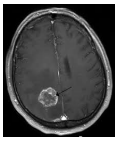
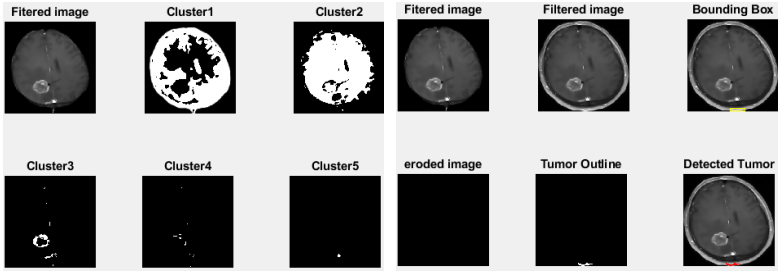
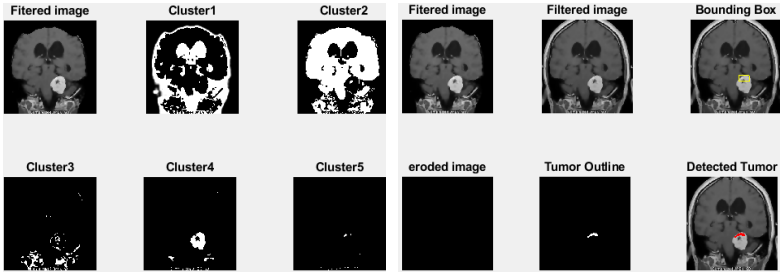
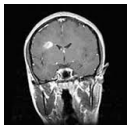
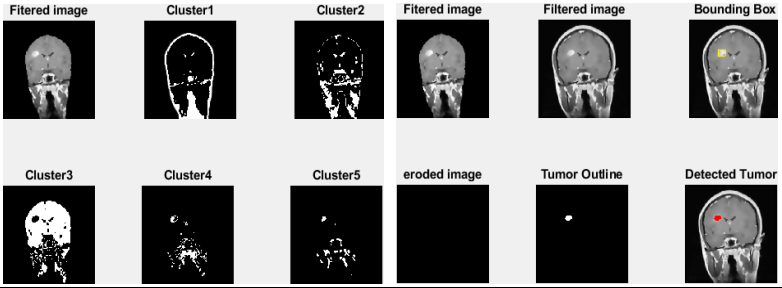
4. Evaluation Results

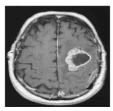
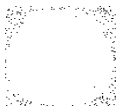
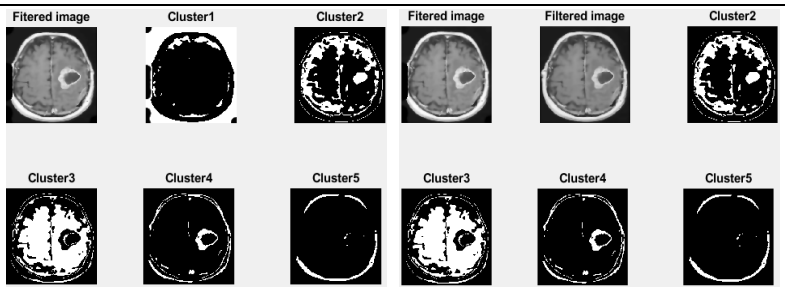
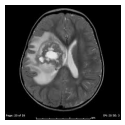
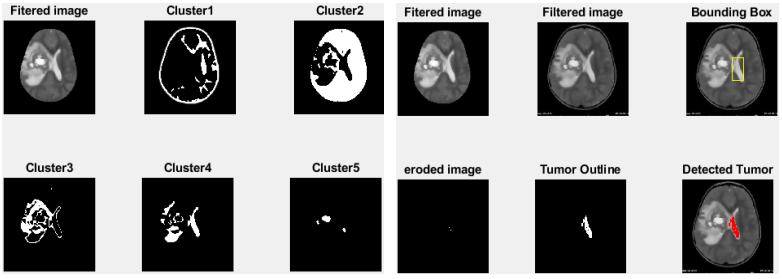
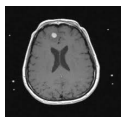
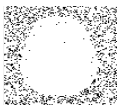
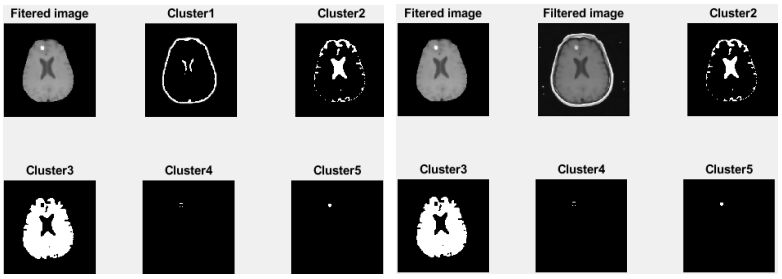
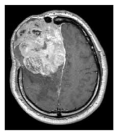
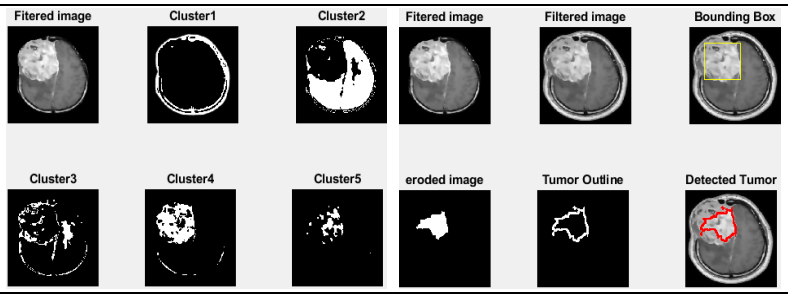
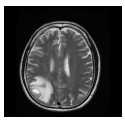
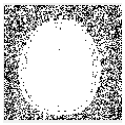
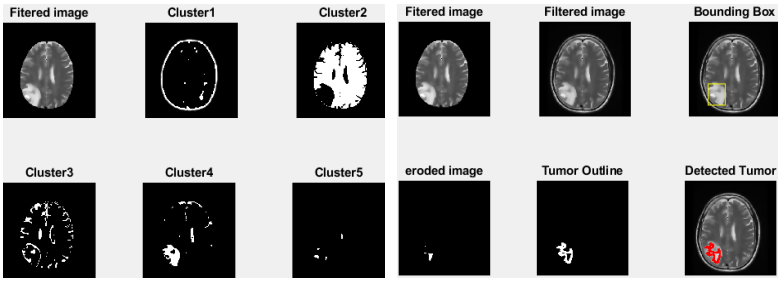
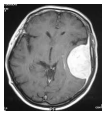
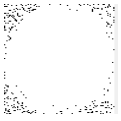
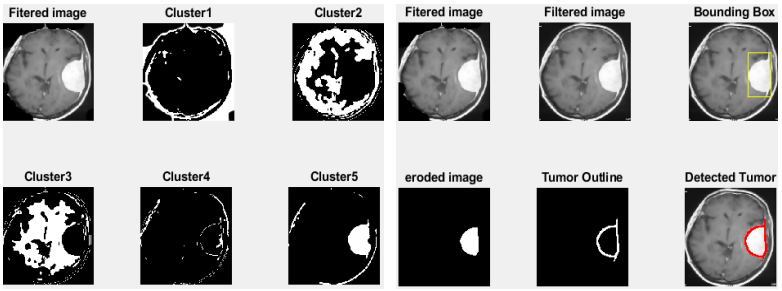
We randomly select 250 images from the kaggle database for the purpose of testing the proposed method [34]. These images are collected from different patients to check the reliability of the proposed method. Evaluation results of 10 images are given in table 1.

5. Result Analysis

The result is assessed on various parameters such as mean, standard deviation, Entropy, RMS, correlation, variance, smoothness, kurtosis, and skewness in order to inspect various aspects of this proposed method.

Table 1. Evaluation results of 10 different images containing benign and malignant tumor

Input Image	Noisy Image	Results after Applying Hybrid Method	Types of Tumor
			Benign
			Malignant
			Benign
			Benign

			Malignant
			Malignant
			Malignant
			Malignant
			Malignant
			Malignant

5.1. Platform for Evaluation

MATLAB implements the proposed model with the specific system requirements as shown in the table 2.

5.2. Result Analysis after Applying Proposed Method

Result analysis based on the statistical parameters is given in table 3. The statistical parameters are mean, standard deviation, entropy, RMS, correlation, variance, smoothness, kurtosis, and skewness. Table 4 predicts the detected tumor area, elapsed time. It is also able to classify the tumor is benign and malignant categories. Tables 5 show the comparative analysis between the size of the image and the elapsed time for brain tumor detection.

5.3. Histogram Analysis of Size of the Image vs. Elapsed Time

Figure 8 shows the histogram analysis of the result shown in table 5. Histogram analysis of the images shown that as size increase the time of brain tumor detection also increases.

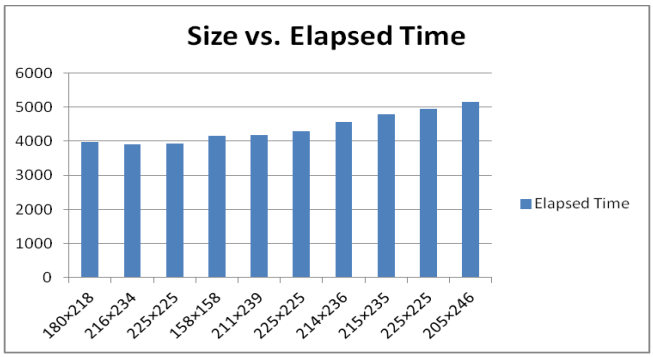


Figure 8: Histogram analysis of 10 different images based on the sizes in pixels and elapsed time in the Matlab

5.4. Comparison Based on the Segmentation Accuracy

Comparative result analysis based on the segmentation accuracy is shown in table 6. The proposed method contains 98% segmentation accuracy.

Table 2. System requirements of the proposed hybrid method

Operating System	Processor	Disk Space	RAM	Graphics
Windows 10	Any Intel or AMD x86-64 processor	2 GB	2 GB	It is suggested to have a hardware accelerated OpenGL 3.3 graphics card with 1 GB GPU RAM.

Table 3. Result evaluation of 10 images with different parameters like mean, standard deviation, entropy, RMS, correlation, variance, smoothness, kurtosis, and skewness

Image	Mean	Standard Deviation	Entropy	RMS	Correlation	Variance	Smoothness	Kurtosis	Skewness
1	0.00324997	0.0944484	3.25081	0.094491	0.149077	0.0089188	0.920936	6.31294	0.346882
2	0.00262226	0.087009	3.31088	0.087039	0.124667	0.0075712	0.921683	7.78956	0.662543
3	0.00311785	0.0857021	3.62058	0.085749	0.091271	0.0073174	0.933302	6.71334	0.514491
4	0.00260813	0.0999868	2.77989	0.100000	0.131878	0.0099708	0.86225	7.30026	0.505607
5	0.00287338	0.0883515	3.65058	0.088388	0.079195	0.0077465	0.927916	5.70919	0.385260
6	0.003997	0.08567	3.21550	0.085750	0.114946	0.007309	0.9472	10.49700	0.829200
7	0.002789	0.08835	3.08390	0.08839	0.106534	0.007804	0.9239	11.15900	0.972100
8	0.002625	0.08701	3.4965	0.08704	0.039998	0.007511	0.9218	6.8506	0.4296
9	0.003545	0.08569	2.6891	0.08575	0.136652	0.007345	0.9409	13.195	1.1142
10	0.004851	0.08968	3.6828	0.0898	0.065291	0.008028	0.9547	5.3377	0.3763

Table 4. Comparative result analysis of 10 images with different parameters like detected tumor area, elapsed Matlab time and type of tumor.

Image	Detected Tumor Area	Elapsed Time(mili seconds)	Types of Tumor
1	753.925	3967.4151	Benign
2	15.9	3898.48409	Malignant
3	56.18	3934.035306	Benign
4	55.385	4159.82001	Benign

5	23.85	4187.05282	Malignant
6	120.575	4283.705	Malignant
7	19.345	4565.229	Malignant
8	563.655	4777.548	Malignant
9	55.385	4952.883	Malignant
10	1444.52	5151.964	Malignant

Table 5. Comparative result analysis based on the size of the image and elapsed time

Image	Size of the Image (Pixels)	Elapsed Time (mili seconds)
1	180×218	3967.4151
2	216×234	3898.48409
3	225×225	3934.035306
4	158×158	4159.82001
5	211×239	4187.05282
6	225×225	4283.705
7	214×236	4565.229
8	215×235	4777.548
9	225×225	4952.883
10	205×246	5151.964

Table 6. Comparative result analysis based on segmentation accuracy

S. No.	Segmentation Technique	Imaging Modalities	Image Task(s)	Segmentation Accuracy
1	Proposed Hybrid Method	MRI	Noise Removal and Segmentation	98%
2	Wavelet transform with morphology based segmentation technique	MRI	Segmentation	80%
3	Computer Aided method that includes the segmentation of the morphological erosion and dilation	MRI	Segmentation	97%
4	Multistage modeling using Eulerian approach based segmentation technique	MRI	Registration and Segmentation	90%
5	Probabilistic neural network based segmentation technique	MRI	Segmentation and Classification	79%
6	Watershed & Thresholding based segmentation technique	CT Images and MRI	Segmentation	92%
7	Geometric transform invariant based segmentation technique	CT Images and MRI	Segmentation	90%
8	Watershed based segmentation technique	MRI	Segmentation	97%
9	SVM and ANN based segmentation technique	MRI	Segmentation and Classification	97.37%

6. Conclusion

The brain images in this study are different in tumor sizes. Therefore, the tumor region is calculated by the proposed process. A variety of statistical and machine learning algorithms such as extended K-mean and modified fuzzy c-means are combined with noise removal adaptive filter technique for implementing such an efficient algorithm. The non-linear morphological erosion and dilation operations carried out to remove the non-tumor portion. As a result, segmentation accuracy and time of brain tumor detection improved with previously discussed techniques in this era. The proposed hybrid method achieves 98% segmentation accuracy. Also, an average of 4 to 5 seconds is the time taken to detect brain tumors depending upon the size of the image. Depending on the outcome review of the various techniques, segmentation level accuracy is higher than other approaches in the background study. It is also found that the statistical parameters such as standard deviation and image are small, which indicates that the statistical analysis outcomes of the proposed method are better than previously discussed algorithms.

Conflict of Interest

So far the knowledge, the authors declare no conflict of interest regarding this article.

Acknowledgment

This work is supported by the director, ASET, Amity University, Lucknow. They would also like to thank the anonymous readers and reviewers for their important and constructive feedback and suggestions, which significantly improved this paper's content.

References

[1] N. Goel, Dr. A. Yadav and Dr. B.M. Singh, “Medical Image Processing: A Review” IEEE International Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity (CIPECH), 57-62, 2016.

[2] Sivakumaran, “Identifying Bone Cancer Using Markov Random Field Segmentation”, 4 June 2018, Available online: https://electronicsforu.com/electronics-projects/prototypes/cancer-markov-random-field-segmentation . [Accessed: 21-August-2019]

[3] Huo, Jing & Okada, Kazunori & van Rikxoort, Eva & Kim, Grace Hyun & R Alger, Jeffry & B Pope, Whitney & Goldin, Jonathan & S Brown, Matthew, “Hybrid segmentation for GBM brain tumors on MR images using confidence-based averaging”. Medical physics, 2013. 40. 093502. 10.1118/1.4817475.

[4] A. Shenbagarajan, V. Ramalingam, C. Balasubramanian and S. Palanivel, “Tumor Diagnosis in MRI Brain Image using ACM Segmentation and ANN-LM Classification Techniques”, Indian Journal of Science and Technology, Vol 9(1), 1-12, Jan 2016.

[5] Aiman Badawi and Muhammad Bilal, ”High-Level Synthesis of Online K-Means Clustering Hardware for a Real-Time Image Processing Pipeline”, J. Imaging 2019, 5, 38; doi:10.3390/jimaging5030038

[6] Yuheng, S.; Hao, Y. Image Segmentation Algorithms Overview. arXiv, 2017; arXiv:1707.02051.

[7] Cardoso, J.S.; Corte-Real, L., “Toward a generic evaluation of image segmentation”, IEEE Trans. Image Process. 2005, 14, 1773–1782.

[8] Chong Zhang , Xuanjing Shen, Hang Cheng, and Qingji Qian, “Brain Tumor Segmentation Based on Hybrid Clustering and Morphological Operations”, Hindawi, International Journal of Biomedical Imaging, Volume 2019, Article ID 7305832, 11 pages https://doi.org/10.1155/2019/7305832.

[9] N. Nezamoddini-Kachouie and P. Fieguth, “A Gabor based technique for image denoising”, Canadian Conference on Electrical and Computer Engineering, 2005. Saskatoon, Sask., 2005, pp. 980-983. doi: 10.1109/CCECE.2005.1557140

[10] James C.Bezdek, RobertEhrlich, WilliamFull, “FCM: The fuzzy c-means clustering algorithm”, Computers & Geosciences volume 10, issues 2–3, 1984, Pages 191-203.

[11] Nameirakpam Dhanachandra, Khumanthem Manglem, Yambem JinaChanu, “Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm”, Procedia Computer Science volume 54, 2015, pages 764-771.https://doi.org/10.1016/j.procs.2015.06.090

[12] Kale Vaishnaw, Vandana B. Malode, “A Novel Approach based on Average Information Parameters for Investigation and Diagnosis of Lung Cancer using ANN”, Pattern Recognition and Image Analysis april 2018, volume 28, Issue 2, pp 301–309.

[13] Ahmed Kharrat, Mohamed Ben Messaoud, Nacéra BENAMRANE and Mohamed ABID, “Detection of Brain Tumor in Medical Images”, International Conference on Signals, Circuits and Systems, PP 1-6, 2009.

[14] Rajeev Ratan, Sanjay Sharma and S. K. Sharma, “Brain Tumor Detection based on Multi-parameter MRI Image Analysis”, ICGST-GVIP Journal, Volume (9), Issue (III), PP- 9- 17, June 2009.

[15] Qurat-Ul-Ain, Ghazanfar Latif, Sidra Batool Kazmi, M. Arfan Jaffar and Anwar M. Mirza, “Classification and Segmentation of Brain Tumor using Texture Analysis”, Recent Advances in Artificial Intelligence, Knowledge Engineering and Data Bases, PP 147- 155, Jan 2010.

[16] M. Usman Akram and Anam Usman, “Computer Aided System for Brain Tumor Detection and Segmentation”, International Conference on Computer Networks and Information Technology, PP 299-302, July 2011.

[17] Stefan Bauer, Christian May, Dimitra Dionysiou, Georgios Stamatakos, Philippe B¨uchler, and Mauricio Reyes, “Multiscale Modeling for Image Analysis of Brain Tumor Studies”, IEEE Transactions on Biomedical Engineering, Vol. 59, No. 1, ,PP 25-29, January 2012.

[18] Arun Tom, P. Jidesh, “Geometric transform invariant Brain-MR Image Analysis for Tumor detection”, IEEE conference on Circuits, Controls and Communications (CCUBE), 1-6, Dec 2013.

[19] R. Vijayarajan and S. Muttan, “Fuzzy C-Means Clustering Based Principal Component Averaging Fusion”, International Journal of Fuzzy Systems, Vol. 16, No. 2, June 2014 PP-153-159.

[20] Meiyan Huang, Wei Yang, Yao Wu, Jun Jiang, Wufan Chen and Qianjin Feng , “Brain Tumor Segmentation Based on Local Independent Projection-based Classification”, IEEE Transactions on Biomedical Engineering , 2633 - 2645 , Vol. 61, Issue 10, Oct 2014.

[21] K. K. Gupta, N. Dhanda and U. Kumar, “A Comparative Study of Medical Image Segmentation Techniques for Brain Tumor Detection”, 2018 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2018, pp. 1-4.

[22] M. Sornam , Muthu Subash Kavitha , R. Shalini, “Segmentation and Classification of Brain Tumor using Wavelet and Zernike based features on MRI”, IEEE International Conference on Advances in Computer Applications (ICACA), Coimbatore, 166-169, Oct 2016.

[23] Rasel Ahmmed, Anirban Sen Swakshar, Md. Foisal Hossain, and Md. Abdur Rafiq, “Classification of Tumors and It Stages in Brain MRI Using Support Vector Machine and Artificial Neural Network”, In Proc. IEEE International Conference on Electrical, Computer and Communication Engineering (ECCE), Bangladesh, 229-237, Feb 2017.

[24] Xiaole Ma, Shaohai Hu and Shuaiqi Liu, “SAR Image De-Noising Based on Shift Invariant K-SVD and Guided Filter”, remote sensing, November 2017.

[25] Arbaz Mukaram, Chidananda Murthy and M.Z.Kurian, “An Automatic Brain Tumour Detection, Segmentation and Classification Using MRI Image”, International Journal of Electronics, Electrical and Computational System, Volume 6, Issue 5 May 2017.

[26] Hanafy M. Ali, “MRI Medical Image Denoising by Fundamental Filters”, in High-Resolution Neuroimaging - Basic Physical Principles and Clinical Applications, Ahmet Mesrur Halefoğlu, Intechopen, pp-111-124, March, 2018.

[27] Saman Riaz, Ali Arshad and Licheng Jiao, "Fuzzy Rough C-Mean Based Unsupervised CNN Clustering for Large-Scale Image Data", applied sciences October 8(10), 1869 , 2018.

[28] Alexander Zotin, Konstantin Simonov, Mikhail Kurako,Yousif Hamad, Svetlana Kirillova, "Edge detection in MRI brain tumor images based on fuzzy C-means clustering", 22nd International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, PP-1262-1270, 2018.

[29] Gupta K.K., Dhanda N., Kumar U., "Depth Analysis of Different Medical Image Segmentation Techniques for Brain Tumor Detection", in: Jain L., Virvou M., Piuri V., Balas V. (eds) Advances in Bioinformatics, Multimedia, and Electronics Circuits and Signals. Advances in Intelligent Systems and Computing, vol 1064, 2020. Springer, Singapore.

[30] Zhe Zhang and Jianhua Song, " A Robust Brain MRI Segmentation and Bias Field Correction Method Integrating Local Contextual Information into a Clustering Model", Recent Advances on Signal Processing and Deep Learning for Public Security and Engineering Applications, Appl. Sci. , 9(7), 1332, March, 2019.

[31] Fangfang Han , Bin Liu , Junchao Zhu and Baofeng Zhang, "Algorithm Design for Edge Detection of High-Speed Moving Target Image under Noisy Environment", Sensors 2019, 19(2), 343; https://doi.org/10.3390/s19020343

[32] Alamgir Nyma, Myeongsu Kang, Yung-Keun Kwon, Cheol-Hong Kim and Jong-Myon Kim, "A Hybrid Technique for Medical Image Segmentation", Hindawi Publishing Corporation, Journal of Biomedicine and Biotechnology, Volume 2012, Article ID 830252, 7 pages, doi:10.1155/2012/830252.

[33] Mohamed Zaki Abderrezak, Mouatez billah Chibane, Prof. Karim Mansour, "A New Hybrid Method for the Segmentation of the Brain MRIs", Signal & Image Processing: An International Journal (SIPIJ) Vol.5, No.4, 77-84, August 2014.

[34] Navoneel Chakrabarty, "Brain MRI Images for Brain Tumor Detection Dataset", Apr, 2019. Available online: https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection. [Accessed: 21-August-2019]

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 6, 19-31 (2019)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

An ML-optimized dRRM Solution for IEEE 802.11 Enterprise Wlan Networks

Mehdi Guessous*, Lahbib Zenkouar

Department of Electrical Engineering, Mohammadia School of Engineers, Mohammed V University In Rabat, Morocco

ARTICLE INFO

Article history:
Received: 09 October, 2019
Accepted: 02 November, 2019
Online: 20 November, 2019

Keywords:
Machine learning
Radio coverage
Radio resources management
Wireless local area networks

ABSTRACT

In an enterprise Wifi network, indoor and dense, co-channel interference is a major issue. Wifi controllers help tackle this problem thanks to radio resource management (RRM). RRM is a fundamental building block of any controller functional architecture. One aim of RRM is to process the radio plan such as to maximize the overall network transmit opportunity. In this work, we present our dynamic RRM (dRRM), WLCx, solution in contrast to other research and vendors' solutions. We build our solution model on a novel per-beam coverage representation approach. The idea of WLCx is to allow more control over the architecture design aspects and recommendations. This dynamization of RRM comes at a price in terms of time and resources consumption. To improve the scalability of our solution, we have introduced a Machine Learning (ML)-based optimization. Our ML-optimized dRRM solution, M-WLCx, achieves almost 79.77% time reduction in comparison with the basic WLCx solution.

1 Introduction

In an enterprise Wlan network, the controller is the central component of the network architecture. The controller manages all the Wifi access points (APs) and provides their radio configuration: channel and transmit power. The controller plays another important role in Wlan integration to other parts of the enterprise network: Local Area Network (LAN), Wide Area Network (WAN), internet, and Datacenter Network (DCN), where application servers reside.

Processing the radio plan is the task of RRM functional architecture block of the controller. It helps minimize co-channel interference and efficient use by APs of the spectrum, thus, optimizing the latter transmit opportunity. Then, how does RRM decide on what channel an access point should use, and at what transmit power?

To build an efficient radio plan that maximizes the network capacity, the controller needs data from APs, Wifi clients or devices (WDs), wired network devices, and servers. This data is what pertains to the quality of the radio interface and client overall experience when accessing the services. However, this information is not sufficient to hint on the whole coverage quality such as the interference at any point in the coverage area. It is only limited to some coverage points, APs and WDs, that are able to monitor the radio interface and report real radio measurements.

To overcome this limitation, either we place sensors everywhere, which is not feasible in an enterprise network (economically and technologically), or model the coverage area. The modelization effort could be done in a laboratory context, by vendors for example, to provide strict recommendations that customers may follow to build their networks. This approach works in common situations. But it requires a lot of engineering effort and monitoring to maintain the network at an optimal condition. In some situations, it may just not work or false the transmit opportunity estimation. For the rest of this work, this approach is referred to as static RRM (sRRM). The third alternative is to allow the controller to do more complex real-time processing without any or very few preconfigured settings and find out the suitable RRM configuration to apply. This approach is the focus of this study and will be referenced as dynamic RRM, or dRRM.

A controller, that supports dRRM, does not rely on preconfigured settings in hardware or software to decide on how to modify the radio plan to meet the utility function. In dRRM, even the system parameters are processed to optimize the network capacity, which is different from sRRM. However, the advantage of dRRM comes at a high price in terms of time,

*Corresponding Author: Mehdi Guessous, Address: Mohammadia School of Engineers, Avenue Ibn Sina, B.P 765, Agdal Rabat 10090, Morocco, Contact No: +212 5377-71905 & Email: mehdiguessous@research.emi.ac.ma

system resources consumption to process the whole network coverage and, adaptation to frequent changes. In this work, we present our dRRM optimization solution that builds on concepts from the Machine Learning (ML) field. Our solution is built on a novel and realistic per-Beam coverage representation approach that is different from related-work encountered research approaches that we discuss later.

In Section 2, we present how related work, research and vendors, process RRM. In Section 3, we present our dRRM solution (that is not sRRM) and compare it to the vendor solution in processing the radio coverage. Before we state the problem in Section 5, and the presentation of our solution in Section 6, we introduce in Section 4, some important facts about Wlan network design, coverage representation models and important machine learning concepts. Section 7 is dedicated to the evaluation of our optimization, before we conclude.

This paper is an extension of the work originally presented in the 2018 15th International Conference on Electrical Engineering, Computer Science and Informatics [1].

2 RRM Related Work

In this section, we discuss RRM approaches from research and vendors of the Wifi market such as Cisco, Aruba-HPE, etc. as they pertain to enterprise Wlan networks. We are interested in algorithms that operate the APs transmit power to maximize the network capacity or optimize radio resource usage. An algorithm is different from another when the used variables are different.

For simplification, we discuss a mono channel condition. This work could then, be easily extended to a multi-channel condition.

2.1 In Research

The first category of approaches concentrate on lower-layer constraints: co-channel interference, physical interface and MAC performance.

The authors in these works [2, 3, 4], modeled the coverage area per-range: transmit, interference, and not-talk ranges, using a circular or disk pattern. The way this model represents the coverage is common but may not hint on some opportunities to transmit as discussed in this work [5].

The author in this work [6], focused instead on the interaction that an AP may have with its neighboring AP. The result is a per-zone, Voronoi zone, negotiated coverage pattern. This model is difficult to put into practice technologically and economically as it was discussed in [5] and [7]. Both models: per-zone and per-range, do not consider upper layer constraints.

Another set of similar works tackle the issue from a power saving perspective. The authors in [8] build their on-demand Wlan approach on the observation of idle APs that have no clients associated. The Wlan controller manages the activation or not of an AP.

The second category of approaches tackles the issue from an upper-layer perspective for applications such as FTP, HTTP. This work [9], as an example, presented an interesting idea to find out a suitable power, or RRM, scheme, that may optimize the application performance. It is a per-experience approach that requires a huge amount of data, to be put into practice. In addition, it is very dependent on the coexistent individual application's behavior. Another challenge is to be able to determine when the physical layer is responsible for the observed performance rather than the application one. Works like [10] use concepts from the Game theory, a powerful tool, to model the interactions between APs. These concepts are applied to the user perception of the QoS it receives. The same limitation of the previously cited work applies to this one also.

The third category tackles the problem from an inter-protocol cooperation point of view like in this example [11]. Making the protocols aware of each other is a good strategy to find an optimum inter-protocol negotiated power scheme that optimizes the performance of each of them individually. It is an idealistic scheme, difficult to put into practice technologically and economically, concerning vendors offering. Let us imagine the integration of a Wifi and a Bluetooth network. The impact of a Wifi AP on a Bluetooth piconet is very important but not the opposite. Then, as an example, it is necessary to find out a way to provide the network controller (for both Wifi and Bluetooth) with the necessary feedbacks so it can adjust the Wifi network power plan to allow a Bluetooth network optimum operation. This would require important data transfers (and power consumption) from the Bluetooth network to the controller, which is very difficult to implement, by design of the Bluetooth devices..

2.2 Vendor Solutions

The approach or theoretical background, behind the vendors' implementations, is hidden in general for commercial purposes; they only provide the settings (recommendations).

Cisco Transmit Power Control (TPC) algorithm, that is a part of Cisco RRM, processes, at each AP, the desired transmit power hysteresis, $Tx_{Hysteresis,Current}$, that is equal to the sum of the current transmit power (initially at maximum), $Tx_{Current}$, and the difference between the power threshold, Tx_{Thresh}, and RSSI^{3rd}, the third neighbor reported RSSI. If the difference between the processed power and the current one, $Tx_{Hysteresis_Thresh}$, is at least $6dBm$, then the current power must be reduced by $3db$ (by half). We should then wait for 10 minutes before re-attempting another calculation. Details about this implementation are given in [12].

Aruba-HPE adopts another strategy. The Adaptive Radio Management (ARM) algorithm maintains two measures for every channel: a coverage index, cov_{idx}, and an interference index, $ifer_{idx}$. The decision of increasing or decreasing the transmit power level on a given channel is based on the processed coverage index as compared to the "ideal" coverage index, noted $cov_{idx,ideal}$, and "acceptable" coverage index, $cov_{idx,acceptable}$, for instance. As a general rule, the current

coverage index should be greater than $cov_{idx,acceptable}$ and equivalent to $cov_{idx,ideal}$. Coverage index, cov_{idx}, corresponds to the sum of two variables : x and y. x is the weighted average of all other APs SNR as being measured by the current AP. y is the weighted average of the processed x variables by other APs from the same vendor and on the same channel. The same thing applies to $ifer_{idx}$ processing. Details of this calculation are in [13].

Fortinet Auto Power that is a part of ARRP, *Automatic Radio Resource Provisioning*, solution, works by reducing automatically the transmit power if the transmit channel is not clear. From the corresponding documentation [14], it is an alternative to manually limiting the number of neighbors per channel (less than 20) by adjusting the transmit power level.

3 Our Dynamic RRM Solution

Our WLCx dynamic RRM solution is based on the per-Beam coverage representation we discuss in the upcoming section. Our solution is "dynamic" because even the parameters' values change: the optimum number of the supported directions per AP, in the case of the WLC2 variant of our solution, as an example. The workflow in Figure 1, describes how our solution works.

Our solution runs three algorithms: TDD (Discovery), TDM (Map) and TDO (Opportunity). After initialization, TDD optimizes the number of supported directions per AP by reducing the power level and doubling the initial number of directions until all neighbors are discovered and at almost one neighbor is discovered per AP direction. Based on information from TDD, TDM categorizes the coverage area points into categories that hints on how these points appear on APs directions. Each category is assigned a cost to hint on its probability to get a fair transmit opportunity. The TDO processes each coverage point opportunity to transmit, taking into account data from TDM and SLA (upper-layer input).

We simulate, using Matlab 2019a, two variants of our WLCx solution: WLC1 and WLC2. In WLC1, all APs share the same optimal number of supported directions and transmit at the same power level. In WLC2, the APs process the same optimal number of the supported directions but may use different transmit power levels per AP. In the same simulation, we compare both WLCx variants to vendor implementation: Cisco. We evaluate models based on their performance at processing the coverage and time this processing takes.

The coverage processing performance, $Pr()$, of a given model, m, is calculated in (1). $I()$, $H()$ and $O()$ are the model processed interference, number of coverage holes and transmit opportunity, respectively.

$$Pr(m) = K_1 \frac{\Sigma I}{I(m)+1} + K_2 \frac{\Sigma H}{H(m)+1} + K_3 \frac{O(m)}{\Sigma O} \qquad (1)$$

The performance calculation in (1), is the weighted sum of relative interference, opportunity and coverage holes in each model. The weights K_1, K_2 and K_3, hints on how important is the processing of interference, opportunity or holes, to the performance of a given model. For the rest of our study, we consider that all variables are of equal importance then, $K_1 = K_2 = K_3 = 1$.

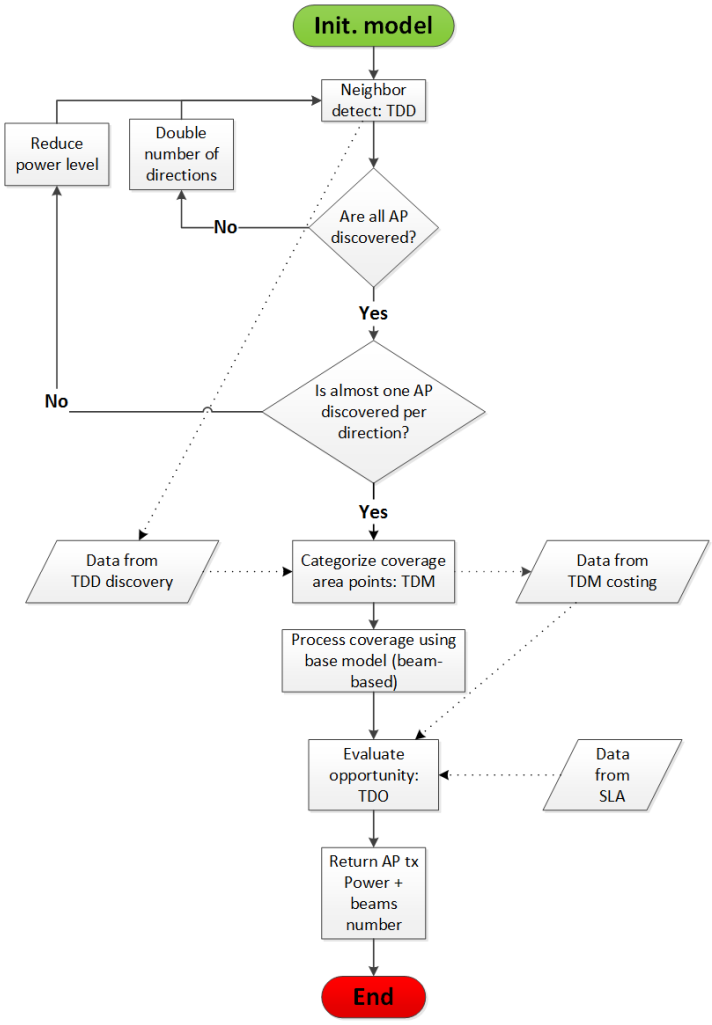


Figure 1: Our WLC dRRM solution workflow

The diagram in Figure 2, shows the performance of models after 10 iterations of the same simulation. Each simulation corresponds to a random distribution of a set of 30 APs and 100 WDs. We check that our WLC2 solution variant performs better than Cisco and WLC1. The Cisco model performance is comparable to WLC1. The processing time of models is represented in Figure 3. The models have a comparable processing time for a large number of the same simulation iterations.

In work [5], we discuss our WLC2 dRRM solution. For further details about our solution, refer to [7] work that is an extension of the previous one.

4 Theoretical Background

Before we dive into the description of the problem, let us recall some facts about Wlan enterprise network architecture design, the importance of coverage representation for

radio planning, and NURBS surfaces concepts that are the foundation of our NURBS optimized WLCx dRRM solution.

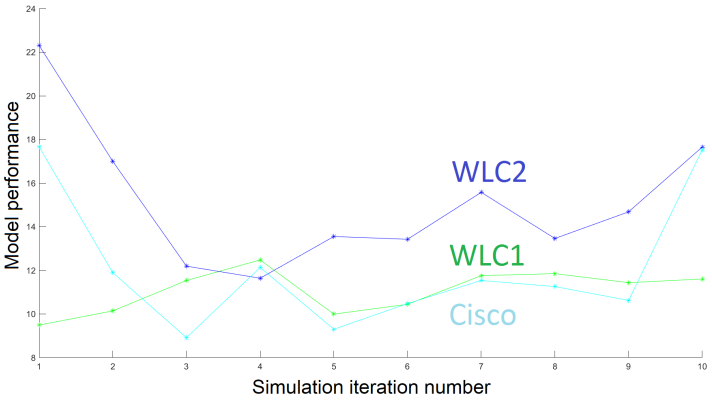


Figure 2: Performance of models after 10 simulations of the network of 30 APs and 100 WDs.

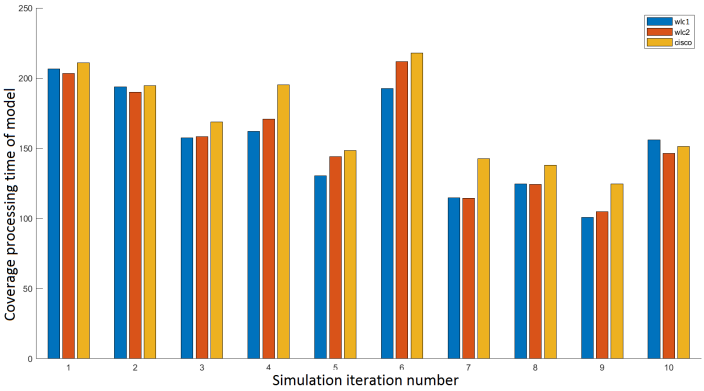


Figure 3: Processing time of models after 10 simulations of the network of 30 APs and 100 WDs.

4.1 *Wifi Unified Architecture*

In a standalone AP-based Wifi architecture, the network capacity does not scale with dense and frequently changing radio environments. To optimize the network capacity, some kind of coordination and control, distributed or centralized, is needed. In UWA, *Unified Wifi Architecture*, a WLC, *Wireless LAN Controller*, acts as a repository of APs intelligence, runs routines to plan radio usage, provides an interface to wired network, etc. and guarantees conformance to policies: QoS and Security, domain-wide, including LAN, MAN, WAN, and DCN, network parts. A typical enterprise Wlan architecture is given in Figure 4. Two market-leading implementations of such WLCs are the Cisco 8540 Wireless Controller and the Aruba 7280 Mobility Controller. The rest of our study focuses on Cisco implementation.

In Figure 4, the APs are located nearest to Wifi clients, WDs. All APs are connected to the LAN and are associated, via Virtual Private Networks (VPN), or tunnels, to the controller, WLC, located at the Datacenter, in a Hub and Spoke architecture. Depending on the network size and requirements, the controller may be located at the same location as the APs. To build an association, an AP should be able to join the controller, via MAN, WAN or internet. After the AP's successful association to the controller, the WDs start their association process that includes authentication, to the corresponding Wlan. Then the WDs access the network resources behind the controller or in some configurations, behind the APs (in FlexConnect or Local Switched mode).

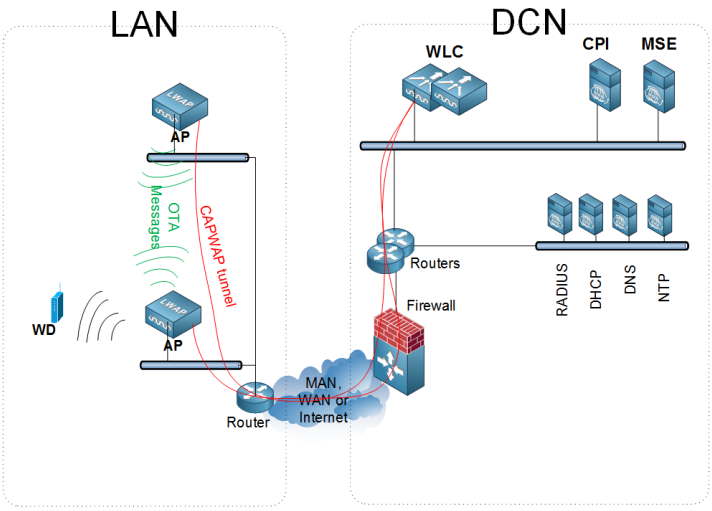


Figure 4: A Wifi unified architecture example topology

WLC receives information about the network from three sources: the wired path toward the datacenter, the radio interface counters of each associated AP, and OTA, *Over-The-Air*, AP-to-AP wireless messages over a dedicated low speed radio. In the case of Cisco, two protocols are available for exchanging data between APs, and between APs and WLC:

- Control and Provisioning of Wireless Access Points (CAPWAP) protocol is used by the APs to build associations to the RF group leader WLC and for control information and data exchange.
- Neighbor Discovery Protocol (NDP) allows the APs to exchange Over-The-Air (OTA) messages that carry standard, per-vendor proprietary control, and management information.

In addition to these protocols, Cisco APs have on-chip features such as CLIENTLINK and CLEANAIR. CLEANAIR enables the APs to measure real-time radio characteristics and send them to the controller via the already established CAPWAP tunnels. Cisco appliances such as Cisco Prime Infrastructure (CPI) and Mobility Services Engine (MSE), shown in Figure 4, extend the capability of this feature to process analytics on Wifi client presence, interfering devices management and heatmaps processing. CLIENTLINK version 4.0, is the Cisco at AP-level implementation of MU-MIMO IEEE 802.11ac beamforming. It works independently of CLEANAIR after the assessment of the quality of the channel. In this scheme, an AP sends a special sounding signal to all its associated WDs, which report, back to this AP, their signal

measurement. Based on these feedbacks, the AP, and not the controller, decides on how much steering toward a specific WD is needed to optimize the energy radiation.

4.2 Radio Coverage Representation Models

We categorize the related-work coverage representation models into three categories: Range-based, Zone-based and Beam-based. In the upcoming subsections, we describe each of them and discuss their limitations.

In the Range-based category of models, it is common to represent an AP's wireless coverage such as: a transmission, interference or no-talk range. These ranges processing is based on the estimation of the distance between the AP and a receiving point P (AP or WD). Further, this category of coverage representation models, consider that an AP's coverage pattern is omnidirectional, with the geometric shape of a circle or a disk, centered at the AP, like in Figure 5. In this scheme, the interference, for example, at any given point is approximated by the weighted intersection of all interfering devices patterns at this point.

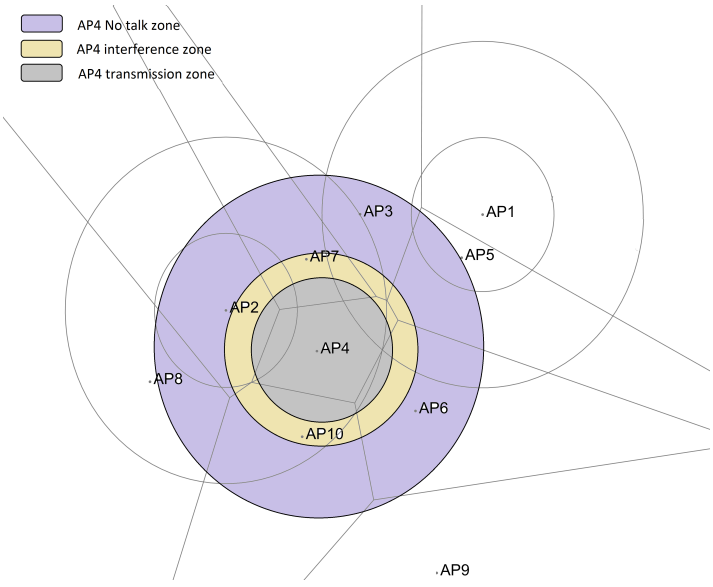


Figure 5: A range-based coverage representation example

In the Zone-based category of models, an AP coverage is a function of its transmission characteristics: channel, power level, etc., but depends also on the neighboring APs. The result of this is that the transmission shape is no more a solid circle but a convex polygon with straight sides. Each straight side defines a borderline that separates two neighboring APs' transmission ranges. The more an AP transmit power is strong, the more the borderline with its neighboring APs is far. Further, it is important to note that a point in a transmission zone of one AP could not be in another AP's transmission zone. An example of Zone-based AP's wireless coverage is represented in Figure 6. In this scheme, the interference caused by the transmission ranges in the previous model, is completely canceled. Only the interference caused by the other ranges: interference, and no-talk ranges, is present.

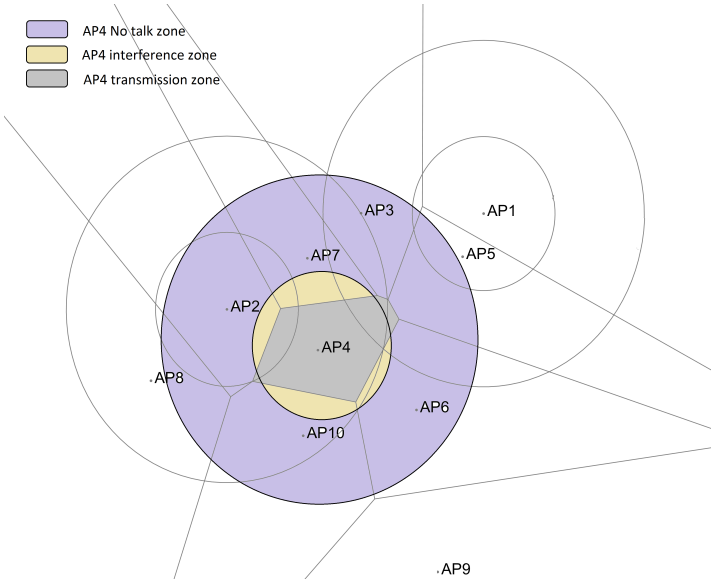


Figure 6: A zone-based coverage representation example

The previous two models: Range and Zone-based, come with these limitations:

- both models are limited to consider that the strength of interference is only inversely proportional to the distance (or quadratic) of an AP from interfering neighbors,
- both models would interpret an increase in a transmission power level as an expanded reach in all directions: uniformly in case of Range-based models but depending on neighboring APs in the case of Zone-based ones,
- a point could not be in two transmission ranges of two different APs at the same time in Zone-based models,
- both models would interpret falsely obstacles to the signal propagation, as a weaker signal from an AP in the context of indoor Wlans does not mean necessarily that this AP is out of reach,
- alternatively, a stronger signal from an AP does not mean necessarily that this AP is at reach: it may be guided or boosted under some conditions.

The consequences of these limitations, the adoption of a Range or Zone-based like representation model of coverage, and regardless of the RRM solution that is built upon, is to false our transmit opportunity processing and misinterpret some phenomena encountered in the specific context of indoor enterprise Wlans.

To overcome the limitations of the previous models, our Beam-based coverage representation, defines for each AP a number of directions over which it may transmit. Depending on the number of directions, their order and transmit power

levels, an AP may be able to mimic a Range or Zone-based scheme. The Figure 7, shows a per-Beam coverage pattern example. In this pattern, the APs have an equal number of directions, equal to eight, that are uniformly distributed and of equivalent transmit power. In works [5, 7], we discussed in detail how per-Zone and per-Range representation models are generalized to per-Beam representation and how our representation model could solve previous models limitations such as : per direction transmit power control, hole coverage reduction, obstacle detection, client localization and transmit opportunities maximization.

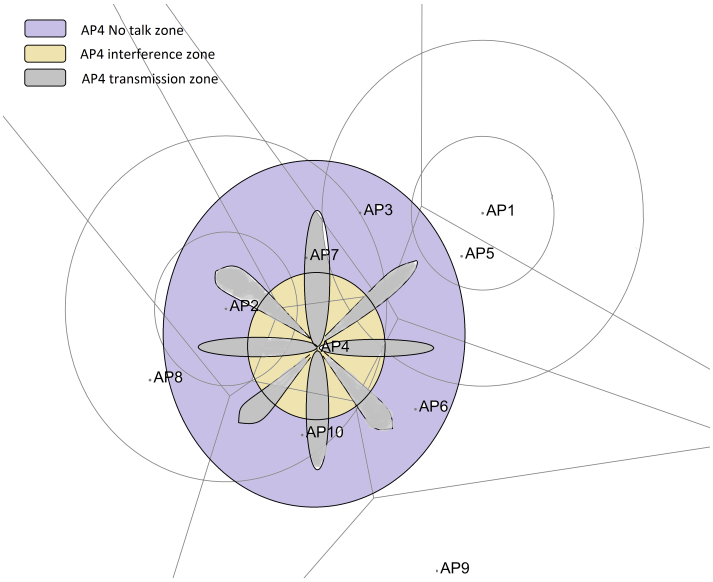


Figure 7: A beam-based coverage representation example

4.3 Machine Learning Regression Models

In his book [15], Tom Mitchell describes a machine, computer program, etc. process of learning when from an experience "E" with respect to some task "T" and some performance measure "P" of this task, its performance on "T", as measured by "P", improves with experience "E".

A simple form of this learning, focus of this work, is described as "supervised" learning. In this learning, the right answers to some input or training data, are provided in advance. Based on this training data, the inputs and corresponding outputs or "truth", the learning algorithm model parameters are processed such as to minimize the error between the predicted outputs and the observed "truth" on the training set.

This "trained" learning, also called hypothesis, is indeed a function that is built using the previously optimized model parameters. This function maps the input variables or features to a predicted outcome.

Supervised learning algorithms could be further classified by the nature of the outcome they work on. If the outcome is continuous, then "regression" models are more suitable. For categorical or discrete outcome values, "classification" algorithms are more suitable.

In our study, we work on outcomes that are continuous, and then we focus solely on regression models. Many types of regression models exist including: linear regression models (LR), regression trees (RT), Gaussien process regression models (GPR), support vector machines (SVM), and ensembles of regression decision trees (BDT).

To choose between models, we compare their Root Mean Square Error (RMSE) validation score. In all the simulations of our work, we observe that Coarse Gaussian SVM and BDT score the best RMSE scores. For the rest of our study, we focus solely on these two models.

Furthermore, SVM and BDT methods represent two distinct algorithm general approaches. In the following subsections, we introduce the important differences in these two approaches.

4.3.1 Support Vector Machines

In SVM, the samples are separated into categories. The idea of this algorithm is to find the maximum gap between these categories. The samples that help find this separation are called support vectors. Each support vector is seen as a dimensional data and the goal of the algorithm is to find the best hyperplane in terms of margin that separates these vectors.

In this work, we use SVM to resolve a linear regression problem. We use a variable that controls the trade-off between the classification errors and the size of the margin. This method corresponds to a soft-margin SVM.

4.3.2 Bagged Decision Trees

Bagging decision trees (BDT) is a meta-algorithm that builds on an ensemble of algorithms that run independently from each other. Each algorithm, called bootstrap, obtains a different result. The result of the meta-algorithm corresponds to the average of the bootstraps individual results. In our case, a bootstrap algorithm may correspond to a simple regression.

The initial training set has n elements. BDT generates, from this set, m new subsets of size n' less than the original set size. If n is equal to n' and n is very big, the probability that each subset has unique values from the initial set is almost 63.2%, the other values are duplicated.

An example of such meta-algorithm is Classification And Regression Trees (CART). The tree operates using a metric and by classifying at each stage, the initial set, the set we are predicting the outcome for. This metric is based on Gini impurity that is calculated from the probability of a certain classification. The classes, which are used in this classification, have been identified during the training phase. In the case of a regression, the algorithm introduces the notion of the variance reduction to build the classes and the corresponding metrics.

5 Problem Description

Coverage processing includes the calculation of interference, opportunity and coverage holes, as per our Beam-based representation model, that is a generalization of the previous work models such as Range or Zone-based representation models.

For the problem description, let us define:

P_i — a coverage point.

$L_{j,k}$ — AP_j, number k direction.

C_i — the sensitiviy of point P_i at reception.

$C_{i,1}$ — AP to which P_i is associated, range of transmission.

$C_{j,2}$ — AP_j, interference range.

$C_{j,3}$ — AP_j, no-talk range.

We show in (2), the interference $I_B()$ that is calculated by WLC2, our WLCx dRRM solution variant, using Beam-based representation model. The processed interference by this model at a point P_i, corresponds to the sum of the intersections, $Sc()$, of all APs beam patterns with C_i and their interference and no-talk ranges with $C_{i,1}$ that is the transmission range of AP_i to which the point P_i is associated.

$$\begin{aligned} I_B(P_i(x,y)) &= \alpha_1 \Sigma_j \Sigma_k \beta_{j,k} * Sc(L_{j,k}, C_i) \\ &+ \Sigma_{k=2}^{3} \alpha_k \Sigma_{j \neq i}^{n} \beta_j * Sc(C_{j,k}, C_{i,1}) \end{aligned} \quad (2)$$

For the opportunity calculation, let us define:

$s_{1,i}$ — passive survey result at a coverage point P_i.

$s_{2,i}$ — active survey result at a coverage point P_i.

In (3), we give the opportunity calculated by WLC2 model, using our Beam-based representation model, $O_B()$. The opportunity is inversely proportional to the interference calculation and hints also, on the result of surveys on the active and passive network paths: $s_{1,i}$ and $s_{2,i}$. Passive surveys allow the controller to have statistics and metrics from the network devices and attached interfaces that are on the network path between the client and the server such as the number of transmit errors, number of lost packets, etc. and is generally available via protocols such: SNMP or *Simple Network Management Protocol*. Active surveys instead, construct traffic patterns and simulate actively the traffic between the client and the server, using protocols such as UDP or TCP, and report measurements such as delay, jitter, etc. to the controller.

$$O_B(P_i(x,y)) \sim \frac{s_{1,i} s_{2,i}}{I_B(P_i(x,y))} \quad (3)$$

The last element to include in the coverage processing, is the number of the detected coverage holes, that is given in (4). Coverage holes are evaluated at every coverage point P_i and correspond to points where the signal is insufficient to perform an accurate communication with their APs of association or the access network, if they are not already associated. $holeThesh_i$ is another variable that is tight to the point P_i sensitivity at reception.

$$H_B(P_i) = |(P_i | \Sigma_j \Sigma_k * Sc(L_{j,k}, C_i) \leq holeThesh_i)| \quad (4)$$

The processing of the coverage, that is done in (2), (3) and (4), is a part of the general processing of our dRRM solution variants: WLC1 and WLC2 that is described in Figure 1 workflow. We give in (5) the necessary time to process the coverage and the changes to this coverage. In (5), we neglected, for simplification, the necessary time to process the optimal number of directions that are supported by the APs and the corresponding transmit power levels. M is the number of APs and any monitoring device. $T_{discovery}$ is the necessary time to run TDD and build a neighborship map. N is the number of coverage points, where the coverage must be calculated. d is the optimum processed number of directions that are supported by APs. $T_{interference}$ corresponds to the necessary time to process coverage. We consider that $T_{interference}$, $T_{opportunity}$ and T_{holes}, times are equivalent.

$$\begin{aligned} T_{WLC2} &= CONST + M * T_{discovery} \\ &+ k(N-M) * M * d * T_{interference} \end{aligned} \quad (5)$$

In Figure 8, we plot the processing time results of models with and without control: simplistic (Range-based), idealistic (Zone-based), WLC1, WLC2 (dRRM) and Cisco (sRRM). We notice that in general, the without-control models perform better than the with-control models due to the addition of the control part of processing. The processing times of the with-control models are equivalent but huge in comparison with the without-control models.

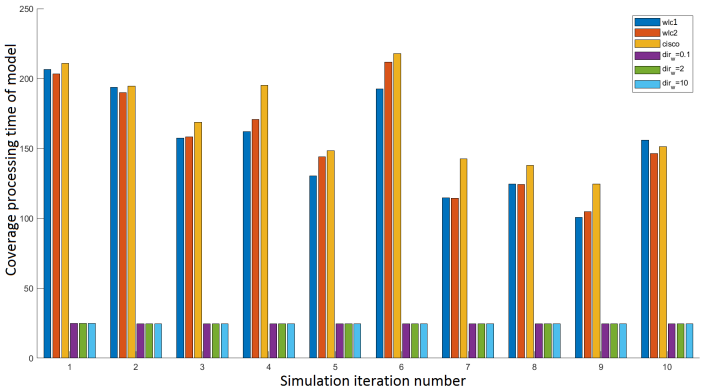


Figure 8: WLC2 time in comparison with idealistic and simplistic models.

The sRRM and dRRM solutions have advantages over each other and over the without-control models, approaches but they require important processing time and resources, which is not suitable in the context of indoor dense enterprise Wlans. In the next section, we propose an optimization solution to with-control RRM models, which is based on concepts from the Machine Learning (ML) field: SVM and BDT. To stick with the aim of this work, we apply this optimization to the example of our dRRM WLC2 solution, but it is easily applicable to the other approaches.

6 MLR-based Optimization Solution: M-WLC2

In this section, we present the details of our solution. We describe the general workflow of our solution, how the training set is chosen and the criteria that allows us to choose between models: RMSE validation score. At the end of this section, we present the new processing time of our solution.

6.1 Workflow

The workflow in Figure 9 describes how our solution processes the coverage. After the initialization, our solution model process a random AP distribution. In this work, we limit the dimension of the coverage plan to 128 coverage points in each dimension. The total number of the coverage points is equal to $128 * 128 = 16,384$ points. From the previous dataset, we build our Test dataset that corresponds to 30% of the total dataset. The Test set is then split into a Train dataset and a Validation dataset. Using the Train dataset, the model trains a number of machine learning models. The predicted data is then compared to the one available in the Validation dataset. Based on the RMSE validation score, the best algorithm is chosen to predict the remaining 70% of the overall dataset.

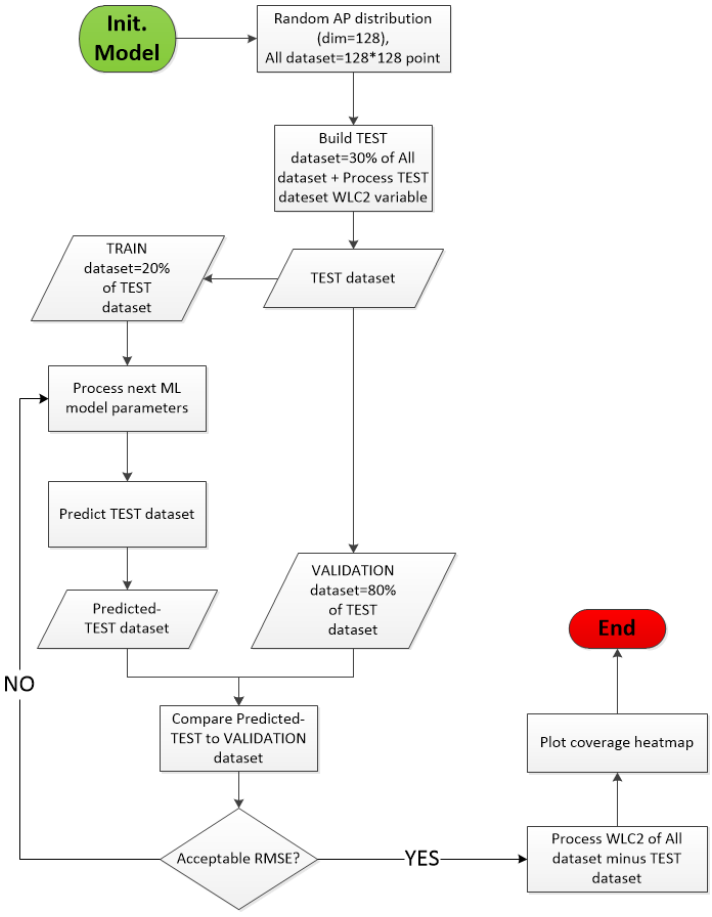


Figure 9: Our ML-WLC2 solution workflow

Our solution is written and simulated in Matlab programming language. Further work would consider the implementation of our solution on Linux-based APs and test its performance in a real laboratory setup. In Figure 10, the detailed description of the software architecture of our solution. The functionProcessML() module is responsible of processing the different data sets, training of the models and of the prediction of the outcomes.

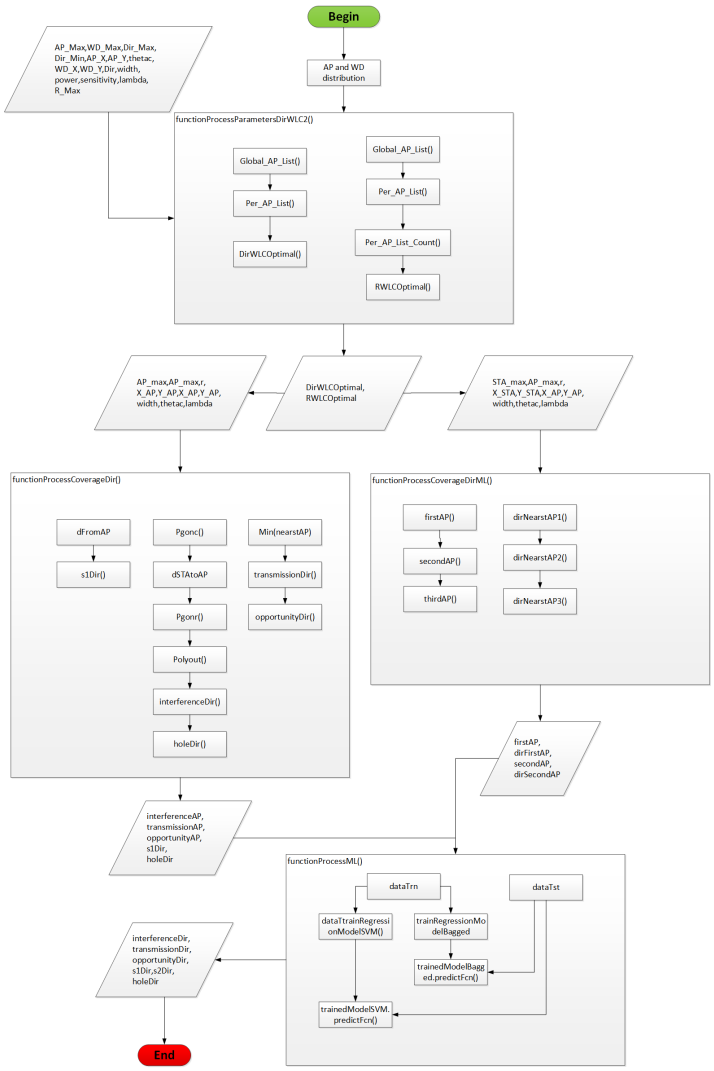


Figure 10: Our ML-WLC2 solution software architecture

6.2 Training Set

To overcome overfitting problems, the training set, Train dataset in the workflow of Figure 9, corresponds to 20% of the Test set. The Validation set is 80% of the the Test set. The Test set represents 30% of the overall available dataset (16,384 points).

For each point in the validation set (the training set is a subset of the validation dataset), we gather this information: the index of the WD or coverage point, the coordinates of this WD, the first AP of association, the corresponding direction of transmission, the transmit power level, the AP load, the

list of neighboring AP, the second AP of association, the corresponding direction of transmission, and so on. The list may be augmented by any relevant information or feature that may have an impact on the phenomenon under study (the interference, in our case).

6.3 RMSE Validation

RMSE stands for Root Mean Square Error. It measures the error between the predicted outcome using the trained model and the truth. The truth in our case is the real measures that were reported by the network APs, WDs, and sniffers.

An important number of tests, the same simulation of the previous networks, allowed us to choose SVM and BDT models. These models score an average of 16.03 and 14.75 points respectively.

Based on the historical results, we set the acceptable RMSE validation score for a model to be accepted in processing the remaining overall dataset points, as it was described in the workflow of Figure 9.

In addition to the RMSE validation score, we use statistical results and visual observations of the heatmap to validate the precision of the outcome.

6.4 Time

The total required processing time of our M-WLC2 solution is given in (6). The total time is the sum of the required time to train the model, T_{train}, and the time to process the validation set, $30\% * T_{WLC2}$, where T_{WLC2} is the total time required to process WLC2 solution model coverage. In case the training set points report the measurement in real-time, this time is very negligible.

$$T_{WLC2,ML} = CONST + 30\% * T_{WLC2} + T_{train} \quad (6)$$

7 Evaluation

In this section, we evaluate the predictions accuracy and the required processing time of the models with and without optimization: WLC2, SVM and BDT M-WLC2.

7.1 Simulation

We simulate a network of 30 APs that are distributed randomly in a 2D plan using the Matlab built-in randperm() function.

The coverage map has a dimension of 128 points in each direction. We process a total of $128 * 128 = 16,384$ coverage points. Each coverage point (in red) corresponds to a potential WD. The WDs are distributed uniformly on the plan.

In Figure 11, we show the distribution of the APs and WDs in the coverage area that corresponds to this simulation. AP1 and AP2 coverage is represented using their Beam-based coverage pattern. AP14 and AP15 coverage is represented by their Range-based coverage pattern.

7.2 Processing of Area Coverage Heatmap

Before we process the coverage of the optimized models, we process the coverage of the WLC2 without optimization. In Figure 12, we show the resultant heatmap of WLC2 processing that indicates the level of interference in the coverage area.

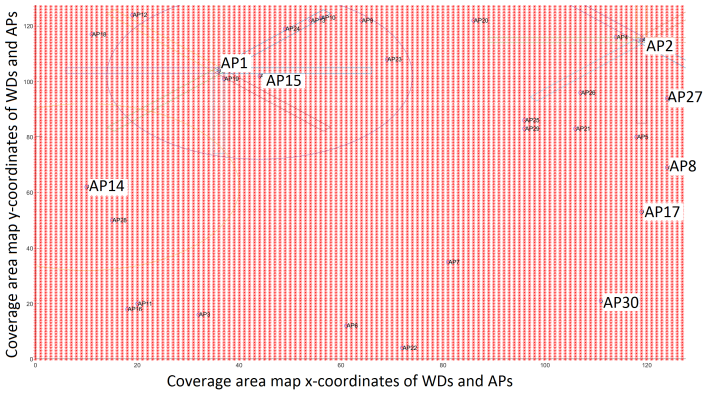


Figure 11: The distribution of APs and WDs in a simulated network

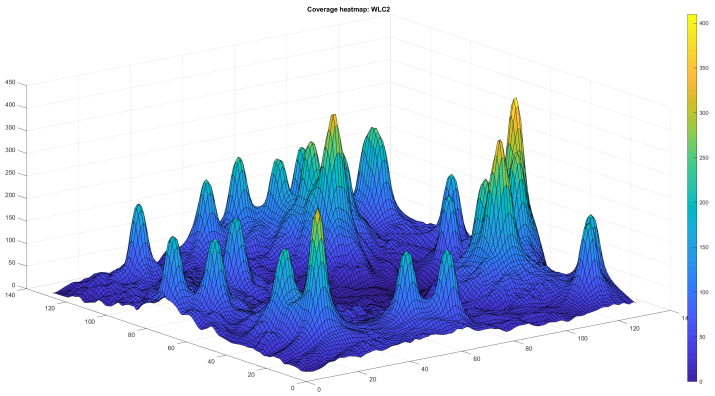


Figure 12: WLC2 area coverage resultant heatmap

In Figure 13 and Figure 14, we show the result of the BDT and SVM optimized WLC2 model coverage processing, respectively. Visually we observe a strong resemblance of the three patterns.

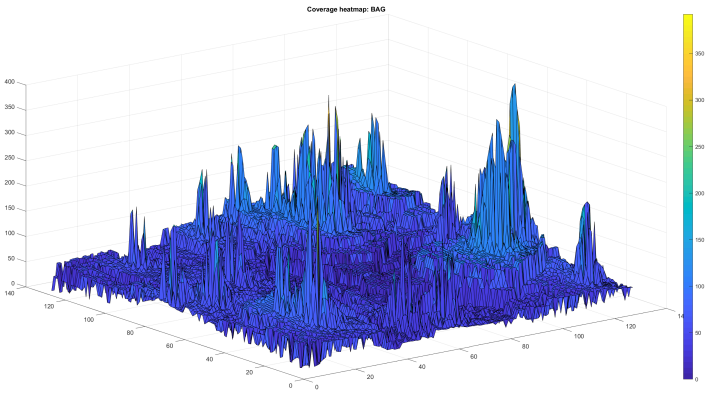


Figure 13: BDT ML-WLC2 area coverage resultant heatmap

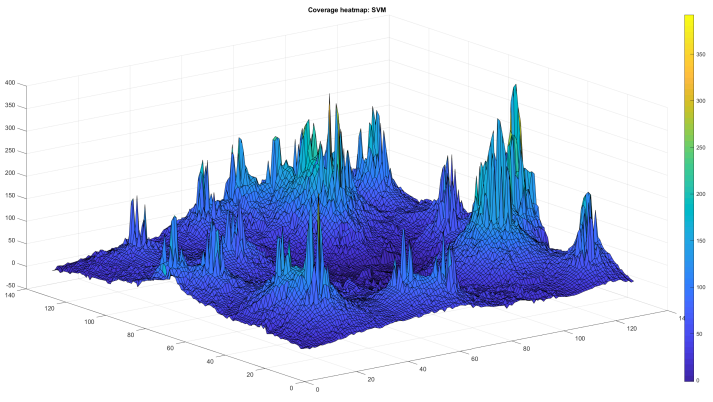


Figure 14: SVM ML-WLC2 area coverage resultant heatmap

In addition to the visual resemblance, the accuracy of the models' prediction is evaluated statistically and using the RMSE validation score. The statistical results hint on the mean, median, standard deviation in between the prediction and the WLC2 model calculated values. The performance of the models depends also on the required processing time.

7.3 *Multiple Iterations of The Same Simulation*

In this section, we present the results of multiple iterations of the same simulation (random distribution of APs at each iteration). In Figure 15, we draw the mean, median and standard deviation results for BDT and SVM models for 10 iterations of the same simulation.

In Figure 15, we observe that BDT presents a better mean around zero and a median slightly around five points in average. SVM shows a better median but a mean around three points in average. BDT standard deviation is twice greater in average than the SVM's, almost 40 points in the first simulations and 30 points in the last ones.

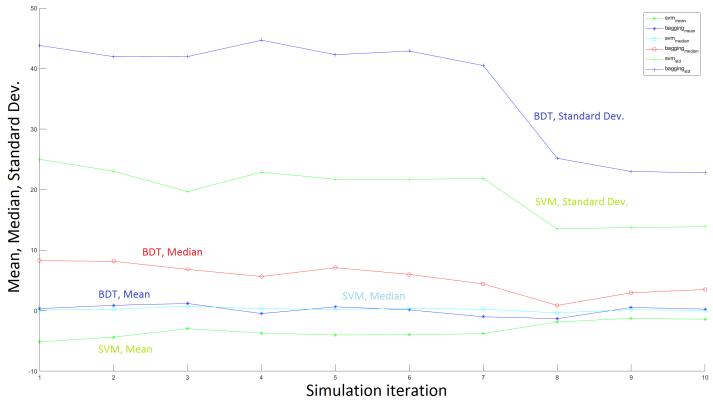


Figure 15: Statistical results of 10 iterations of BDT and SVM M-WLC2 simulation

For the same simulations, we show in Figure 16, the variation of the RMSE score. We observe that the BDT solution scores the best RMSE value, almost 45 points, against the SVM model, almost 25 points, in the first simulations. In the three last simulations, the RMSE score drops to almost 25 points for BDT model and only 15 points for the SVM model.

In Figure 17, we show the required processing time of models with and without optimization. In general, we observe that the optimized models: BDT and SVM times are comparable and very negligible in comparison with WLC2 time, almost 2.5 smaller.

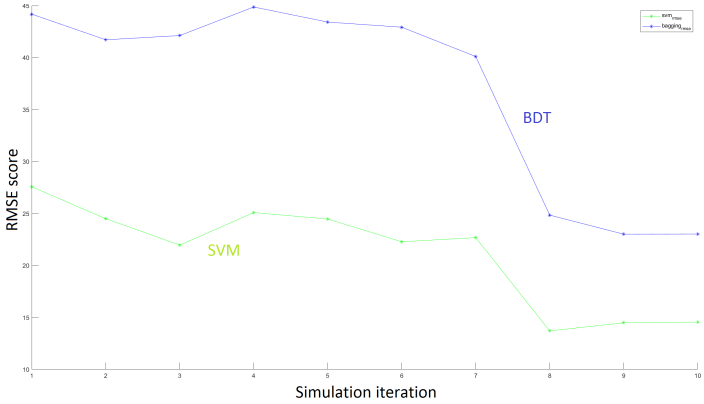


Figure 16: RMSE resultats of 10 iterations of BDT and SVM M-WLC2 simulation

7.4 *The Effect of Modifying The Training Set Size*

In this subsection, we propose to check the effect of modifying the training set size on the optimized models performance.

In Table 1, we show the results of modifying the size of the training set in this range: 10%, 20%, 30%, 50% and 90% of the total available data set. We notice no remarkable change in statistical results in terms of mean and median in between the models when the training set is 10%, 20% and 30%. For bigger training sets, these values are getting remarkably smaller. The SVM model standard deviation decreases with the increasing training set sizes. Concerning the BDT model, we observe the opposite; the standard deviation is at its highest value when the size of the training set is 90% of the total available data set.

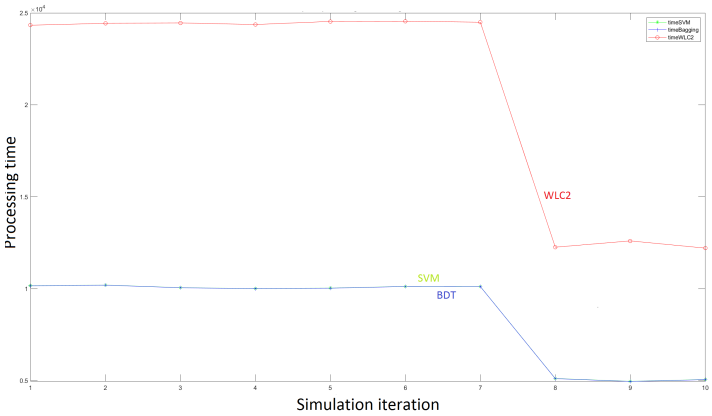


Figure 17: Required processing time of 10 iterations of BDT and SVM M-WLC2 simulation, training set 30% of total data set

Table 1: Effect of modifying the training set size on model performance

Var.	trn_{set} = 10%	20%	30%	50%	90%
m_{svm}	-3,92	-3,23	-3,25	-2,39	-2,50
m_{bdt}	-0,27	0,18	0,11	-0,03	-0,05
med_{svm}	0,29	0,25	0,20	0,15	0,12
med_{bdt}	4,69	5,66	5,37	4,87	5,45
std_{svm}	23,76	20,55	19,69	16,50	16,40
std_{bdt}	34,93	34,75	36,90	35,90	38,91
$rmse_{svm}$	25,10	21,88	21,13	17,82	17,21
$rmse_{bdt}$	34,69	34,74	37,02	36,02	38,85
t_{svm}	2594 ,88	3776 ,93	5291 ,24	7602 ,12	1485 1,02
t_{bdt}	2594 ,19	3774 ,37	5286 ,25	7585 ,18	1479 3,40
t_{wlc2}	1209 1,41	1202 4,33	1280 9,81	1229 2,02	1469 0,10

Table 2: Modifying the dimension of heatmap effect on ML heatmap processing

Var.	dim = 256	128	64	32	16
m_{svm}	-2,15	-4,16	-1,87	-0,93	0,04
m_{bdt}	-0,15	-0,11	0,09	0,34	0,64
med_{svm}	0,15	0,36	0,32	0,19	0,69
med_{bdt}	4,15	5,56	1,59	-0,73	-0,67
std_{svm}	15,82	24,8	19,83	23,76	23,29
std_{bdt}	25	37,6	23,72	20,21	17,47
$rmse_{svm}$	16,04	26,7	20,94	24,08	25,63
$rmse_{bdt}$	24,75	37,63	23,29	21,02	18,2
t_{svm}	1297 1,54	2744 ,41	450,09	130,53	58,9
t_{bdt}	1295 9,84	2743 ,7	450,12	130,65	58,99
t_{wlc2}	6406 3,31	1280 9,94	1836 ,07	304,17	76,58

In general, BDT scores better RMSE values than SVM. SVM's RMSE score is better when the size of the training set is the smallest, 10% of the total available data set. Differently from SVM, the BDT model scores the best RMSE value when the training set is 90% of the total data set.

The optimized models processing time is comparable. For a training set equal to 10% of the total available data set, this time is enhanced by almost 78.54% relatively to the WLC2 model.

7.5 Modifying The Dimension Size of The Coverage Area

In this subsection, we present the results of modifying the size of the coverage area or dimension on the performance of the optimized models. The training set is set to 10% of the total available data set.

In Table 2, we observe that the required time to process a 256 dimension coverage area is almost 842 times higher than processing a 16 dimension map. In the first case, our optimized models enhance the required time by almost 79.77%.

The visual aspect of the coverage is presented in Figure 18. The pattern is comparable between both optimized models and representative of WLC2 result. We notice though that SVM result is more precise than BDT if we focus on the lowest coverage values.

The BDT model scores the best RMSE score in comparison with the SVM models for the high dimensions. The opposite happens for the low dimension values.

In terms of the statistical results, the SVM model results in less standard deviation from the mean. The SVM median is around zero whereas the BDT model median is at almost 5 points for high dimensions.

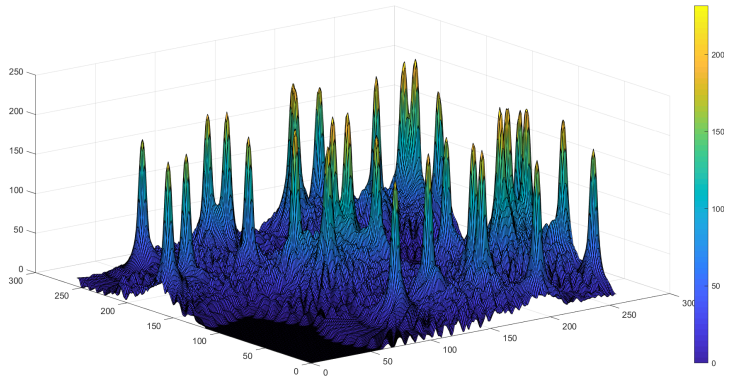
8 Conclusion

In this work, we have introduced our WLC2 dRRM solution in contrast with literature approaches: Zone or Voronoi diagram-based (idealistic) and Range-based (simplistic), and vendor sRRM category of models: Cisco especially. We have shown that our solution performs better than the vendor sRRM solution in a simulated controller-based Wifi environment.

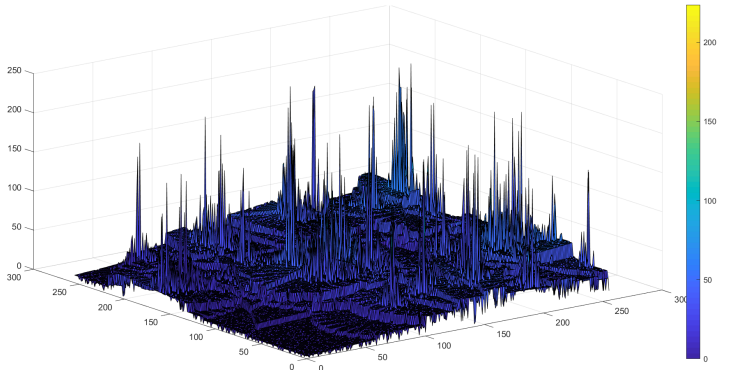
Further, in this work, we have shown that the basic variant of our dRRM solution, WLC2, comes with some limitations: the important system resources consumption and the required processing time. The M-WLC2 optimization solution, which is based on important machine learning concepts, allowed us to achieve an average of 79.99% relative time reduction by processing only 10% of the total available data set. The accuracy of the results was evaluated visually and statistically.

Our M-WLC2 optimization solution does not depend on the Beam-based coverage representation model approach we adopted for the simulation of the coverage area. It relies only on the environmental variables, which influence the phenomena, to build a prediction model rather than on an analytical calculus of the physical phenomena itself. Besides, our optimization solution is not limited to only our dRRM solution but could be easily extended to optimize sRRM models too.

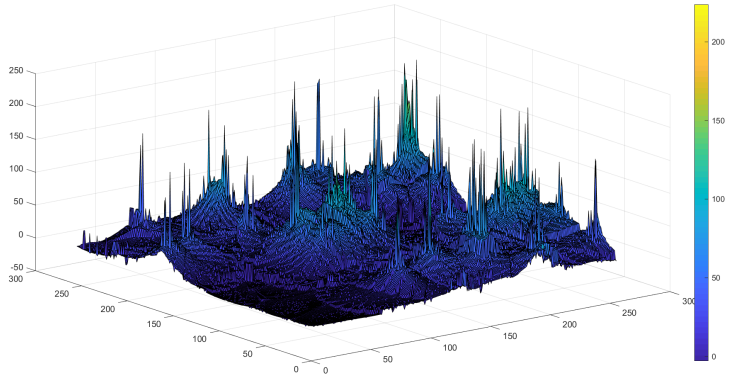
It is to mention that the optimization process of our dRRM solution, presented in this work, could be accomplished in different ways. The first approach is to describe analytically the physical phenomena under study like in our solution N-WLCx discussed in this work [16]. The second approach is to work deeply on the performance of the actual machine learning algorithms as it may be suggested by works such as [17]. The idea here is to adapt the standard machine learning algorithms to our specific need.



(a) "WLC2"



(b) "BDT"



(c) "SVM"

Figure 18: Visual aspect of ML optimized WLC2 coverage when $dim = 256$

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment We would thank colleagues : researchers, engineers, and reviewers for sharing their precious comments and on-field experience that improved the quality of this paper.

References

[1] M. Guessous and L. Zenkouar. "ML-Optimized Beam-based Radio Coverage Processing in IEEE 802.11 WLAN Networks". In: *2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*. Oct. 2018, pp. 564–570. DOI: 10.1109/EECSI.2018.8752874.

[2] Rafiza Ruslan and Tat-Chee Wan. "Cognitive radio-based power adjustment for Wi-Fi". In: *TENCON 2009-2009 IEEE Region 10 Conference*. IEEE. 2009, pp. 1–5. DOI: 10.1109/TENCON.2009.5396078.

[3] Daji Qiao et al. "Adaptive transmit power control in IEEE 802.11 a wireless LANs". In: *The 57th IEEE Semiannual Vehicular Technology Conference, 2003. VTC 2003-Spring.* Vol. 1. IEEE. 2003, pp. 433–437. DOI: 10.1109/VETECS.2003.1207577.

[4] Nabeel Ahmed and Srinivasan Keshav. "A successive refinement approach to wireless infrastructure network deploy-

ment". In: *IEEE Wireless Communications and Networking Conference, 2006. WCNC 2006*. Vol. 1. IEEE. 2006, pp. 511–519.

[5] Mehdi Guessous and Lahbib Zenkouar. "Cognitive directional cost-based transmit power control in IEEE 802.11 WLAN". In: *2017 International Conference on Information Networking (ICOIN)*. IEEE. 2017, pp. 281–287. DOI: 10.1109/ICOIN.2017.7899520.

[6] Prateek R Kapadia and Om P Damani. "Interference-constrained wireless coverage in a protocol model". In: *Proceedings of the 9th ACM international symposium on Modeling analysis and simulation of wireless and mobile systems*. ACM. 2006, pp. 207–211. DOI: 10.1145/1164717.1164754.

[7] Mehdi Guessous and Lahbib Zenkouar. "A novel beamforming based model of coverage and transmission costing in IEEE 802. 11 WLAN networks". In: *Adv. Sci. Technol. Eng. Syst. J.* 2.6 (2017), pp. 28–39. DOI: 10.25046/aj020604.

[8] Amit P Jardosh et al. "Green WLANs: on-demand WLAN infrastructures". In: *Mobile Networks and Applications* 14.6 (2009), pp. 798–814. DOI: 10.1007/s11036-008-0123-8.

[9] Aditya Akella et al. "Self-management in chaotic wireless deployments". In: *Wireless Networks* 13.6 (2007), pp. 737–755. DOI: 10.1007/s11276-006-9852-4.

[10] Cem U Saraydar, Narayan B Mandayam, David J Goodman, et al. "Efficient power control via pricing in wireless data networks". In: *IEEE transactions on Communications* 50.2 (2002), pp. 291–303. DOI: 10.1109/26.983324.

[11] Dipankar Raychaudhuri and Xiangpeng Jing. "A spectrum etiquette protocol for efficient coordination of radio devices in unlicensed bands". In: *14th IEEE Proceedings on Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003.* Vol. 1. IEEE. 2003, pp. 172–176. DOI: 10.1109/PIMRC.2003.1264255.

[12] *Radio Resource Management White Paper - Transmit Power Control (TPC) Algorithm [Cisco 5500 Series Wireless Controllers]*. en. June 2016. URL: https://www.cisco.com/c/en/us/td/docs/wireless/controller/technotes/8-3/b_RRM_White_Paper/b_RRM_White_Paper_chapter_0101.html (visited on 05/14/2019).

[13] *ARM Coverage and Interference Metrics*. URL: https://www.arubanetworks.com/techdocs/ArubaOS_64x_WebHelp/Content/ArubaFrameStyles/ARM/ARM_Metrics.htm (visited on 05/17/2019).

[14] *Lowering the power level to reduce RF interference*. URL: https://help.fortinet.com/fos50hlp/52data/Content/FortiOS/fortigate-best-practices-52/Wireless/Lowering_Power_Level.htm (visited on 05/17/2019).

[15] Thomas M. Mitchell. *Machine Learning*. 1st ed. New York, NY, USA: McGraw-Hill, Inc., 1997. ISBN: 0070428077, 9780070428072.

[16] Mehdi Guessous and Lahbib Zenkouar. "A nurbs based technique for an optimized transmit opportunity map processing in wlan networks". In: *International Conference on Wired/Wireless Internet Communication*. Springer. 2017, pp. 143–154. DOI: 10.1007/978-3-319-61382-6_12.

[17] Alireza Babaei. "Longitudinal vibration responses of axially functionally graded optimized MEMS gyroscope using Rayleigh–Ritz method, determination of discernible patterns and chaotic regimes". In: *SN Applied Sciences* 1.8 (2019), p. 831. DOI: 10.1007/s42452-019-0867-8.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 2, 19-27 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Distributed Linear Summing in Wireless Sensor Networks with Implemented Stopping Criteria

Martin Kenyeres[*,1], Jozef Kenyeres[2]

[1]*Institute of Informatics, Slovak Academy of Sciences, Dubravska cesta 9, 845 07 Bratislava 45, Slovak Republic*

[2]*Sipwise GmbH, Europaring F15, 2345 Brunn am Gebirge, Austria*

ARTICLE INFO

Article history:
Received: 04 January,2020
Accepted: 31 January, 2020
Online: 09 March, 2020

Keywords:
Distributed summing
Average consensus
Wireless sensor networks

ABSTRACT

Many real-life applications based on the wireless sensor networks are equipped with data aggregation mechanisms for suppressing or even overcoming negative environmental effects and data redundancy. In this paper, we present an extended analysis of the linear average consensus algorithm for distributed summing with bounded execution over wireless sensor networks. We compare a centralized and a fully-distributed stopping criterion proposed for the wireless sensor networks with a varied initial configuration over random geometric graphs with 200 vertices in order to identify the optimal initial configuration of both analyzed stopping criteria and the algorithm as well as finding out which stopping criterion ensures higher performance of the examined algorithm in terms of the estimation precision and the convergence rate expressed as the iteration number for the consensus.

1 Introduction

This paper is an extension of work originally presented in the conference Proceedings of the IEEE 17st World Symposium on Applied Machine Intelligence and Informatics, (SAMI 2019) [1].

1.1 Wireless Sensor Networks

Wireless sensor networks (WSNs), a technology operating in an unattended ad-hoc mode, may be formed by hundreds to thousands autonomous sensor nodes[1], example shown in Figure 1, to gather information from the adjacent environment and to detect certain events of interest over the monitored area (e.g., buildings, homes, forests, oceans, mountains, etc.) [2–5]. Over the past years, this technology has found the application in many different areas, e.g., disaster detection, industrial monitoring, health assistance, military surveillance, etc. [6]. Therefore, the design of the sensor nodes has to be adapted to environmental conditions (e.g., the sensor nodes for underground operation have an increased transmission power whereby noisy channel attenuations can be overcome) [3]. However, as these sensor nodes are simple and low-cost devices, they suffer from limited energy capacity and computation capabilities, resulting in decreased robustness to potential threads such as radiations, pressure, temperature, attacks, etc. [7–9]. Thus, many of the real-life WSN-based applications are equipped with data aggregation mechanisms for suppressing or even overcoming the impacts of these negative factors on the operation of WSNs. Moreover, data aggregation can eliminate highly correlated and duplicated information as well [8].

1.2 Data Aggregation

In general, data aggregation poses any process in which data from multiple sources is transformed into a summary form for further statistical processing or analyzing. The data aggregation mechanisms are designed to process measured information from multiple independent sensor nodes in order to provide a greater quality of servise (QoS) [8]. Xiao et al. define two categories of the data aggregation mechanisms, namely [9]:

- *Centralized*
- *Distributed*

In the centralized data aggregation schemes, the sensor nodes transmit their measured information to a fusion center either directly or by multi-hop communication. In the second scenario, each sensor node has to establish and update the table with routing information, which is not a too effective way to deliver information to a fusion center over mobile networks and networks with limited energy sources such as WSNs [9]. Therefore, the distributed schemes

*Corresponding author: Martin Kenyeres, E-mail: martin.kenyeres@savba.sk
[1]see Figure 1 for an example

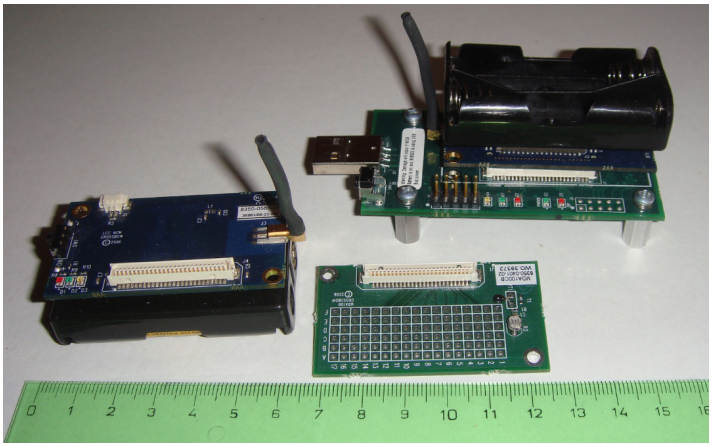


Figure 1: Example of WSNs – Memsic's classroom kit (Source: Reprinted from [13])

are more frequently applied nowadays [10–13]. In these schemes, the adjacent sensor nodes communicate with each other and update their states according to the collected data [14]. Thus, a fusion center is not necessary for data aggregation as well as the sensor nodes know only local information whereby data aggregation is significantly simplified. In [15], Jesus et al. divide distributed aggregation schemes into three groups in terms of the communication perspective:

- *Structured*
- *Unstructured*
- *Hybrid*

Structured algorithms are dependent on network topologies and routing algorithms and therefore are not appropriate for mobile systems. Moreover, a single point of failure (e.g., a dead node, a link failure) can have a fatal impact on a whole system, especially in sparsely connected topologies such as tree-based structures. The next category, *Unstructured algorithms*, is independent of network topology and structure unlike the first category, and thus, network topology does not have to be predefined. These algorithms are characterized by simplicity, high scalability, high robustness, etc. The last category, i.e., *Hybrid algorithms*, utilizes strengths and suppress weaknesses of the previous approaches by their combination.

In terms of the computation perspective, the distributed data aggregation mechanisms can be classified as follows [15]:

- *Computation of decomposable functions*
- *Computation of complex functions*
- *Counting*

As mentioned earlier, we focus our attention on linear consensus-based algorithms for distributed summing, which can be classified as unstructured algorithms for computation of decomposable functions. These algorithms have significantly attracted the attention of the academy recently [16–18]. In [19], Gutierrez-Gutierrez et al. define two categories of these algorithms, namely:

- *Deterministic*
- *Stochastic*

From this set of the algorithms, we choose deterministic linear average consensus (referred to as AC), characterized by a variable mixing parameter ϵ, for distributed summing [20].

A properly configured stopping criterion can significantly optimize data aggregation over WSNs; therefore, its implementation is an essential complement for distributed data aggregation mechanisms. As mentioned above, we assume that the execution of the examined algorithm is bounded by a stopping criterion in our analyzes.

1.3 Summary of Contribution

In this paper, we analyze AC for distributed summing whose execution is bounded by a stopping criterion. We vary its initial configuration and the configuration of the analyzed stopping criteria (we examine a centralized and a fully-distributed one) in order to identify which configurations ensure the highest performance of the examined algorithm and in order to compare the analyzed stopping criteria in terms of the precision and the convergence rate. As mentioned earlier, a properly configured stopping criterion and algorithm optimize data aggregation, thereby saving energy, reducing the computation/the communication requirements, prolonging the network lifetime, etc.

In this paragraph we clarify the novelty of the presented paper in a comparison with the existing work presented in [1]: In [1], we analyze AC bounded by the fully-distributed stopping criterion defined in [13] - we vary the mixing parameter and the parameters of the implemented stopping criterion in order to identify the best-performing initial configurations of both the algorithm and the stopping criterion. In that paper, the inner states are multiplied by the network size before AC begins. In this paper, we compare these results with the scenario when the inner states are multiplied after AC is completed. The goal of this contribution is to identify when it is the most optimal to carry this multiplication out in terms of the precision and the convergence rate expressed as the iteration number for the consensus achievement. Furthermore, we also analyze the performance of the centralized stopping criterion from [6] and compare it with the results from [1] and the results from the

extended analysis presented in this paper in order to identify which stopping criterion achieves higher performance.

1.4 Paper Organizanization

The paper is organized as follows: Section 2 is concerned with theoretical insight into the topic, i.e., we provide a mathematical model of AC and WSNs, the convergence conditions, and an explanation of the analyzed stopping criteria. Section 3 deals with the applied methodology and the used metrics for performance evaluation. Section 4 consists of the experimentally obtained results and a consecutive discussion. Conclusions are presented in Section 5.

2 Theoretical Background

The primary purpose of AC is to estimate the arithmetic mean from all the initial states. However, it can fulfill also other functionalities - in this paper, we analyze AC for distributed summing, i.e., the sum from all the initial states is estimated. In this case, the information about the network size n is necessary to be known by each sensor node, and the inner states have to be multiplied by this value, otherwise, the arithmetic mean is estimated instead of the sum [1].

2.1 Model of AC for Distributed Summing in WSNs

We model WSNs as simple finite graphs determined by the vertex and the edge sets, i.e., $G = (\mathbf{V}, \mathbf{E})$ [20]. The vertex set $\mathbf{V}$ is formed by all the graph vertices, which represent the sensor nodes in a network, i.e., $\mathbf{V} = \{v_1, v_2, ... v_n\}$. The graph order is equal to the number of the sensor nodes in a network, and so, $|\mathbf{V}| = n$. Two vertices v_i and v_j are adjacent when they are connected to one another by an edge, i.e., $e_{ij} \in \mathbf{E}$.

In AC, each vertex v_i stores and update its inner state $x_i(k)$[2], represented by a scalar value - the inner states are initiated by for example a local measurement. After the initialization, the inner states asymptotically converge to the arithmetic mean from all the initial inner states by being updated at each iteration, as described in 1, [21]:

$$\mathbf{x}(k+1) = \mathbf{W} \times \mathbf{x}(k) \tag{1}$$

Here, $\mathbf{x}(k)$ is a variant column vector gathering all the inner states at the corresponding iteration, and $\mathbf{W}$ poses the weight matrix, affecting many aspects of the algorithm such as the convergence rate of the algorithm, the initial configuration, meeting/violating the convergence conditions, etc. [22]. As already stated, the inner states asymptotically converge to the estimated aggregate function, which can be expressed using 2, [21]:

$$\lim_{k\to\infty} \mathbf{x}(k) = \lim_{k\to\infty} \mathbf{W}^k \times \mathbf{x}(0) = \frac{1}{n}.\mathbf{1} \times \mathbf{1}^{\mathrm{T}} \times \mathbf{x}(0) \tag{2}$$

Here, $\mathbf{1}$ represents a column all-ones vector formed by n elements. The algorithm works correctly iff the limit from (2) exists, which is ensured by meeting these convergence conditions described by 3–5, [21], [23, 24]:

$$\mathbf{1}^{\mathrm{T}} \times \mathbf{W} = \mathbf{1}^{\mathrm{T}} \tag{3}$$

$$\mathbf{W} \times \mathbf{1} = \mathbf{1}, \tag{4}$$

$$\rho(\mathbf{W} - \frac{1}{n} \cdot \mathbf{1} \times \mathbf{1}^{\mathrm{T}}) < 1 \tag{5}$$

Here, $\rho(\cdot)$ represents the spectral radius of the corresponding vector/matrix, which is the largest eigenvalue in the modulus, i.e. 6, [22]

$$\rho(\cdot) = \max_i \{| \lambda_i(\cdot) |\} \tag{6}$$

One of the ways to configure the weight matrix $\mathbf{W}$ is to allocate a constant value[3] to all the edges in a graph. Such a matrix is referred to as the Peron Matrix and is defined by 7, [25]:

$$[W]_{ij} = \begin{cases} \epsilon, & \text{if } e_{ij} \in \mathbf{E} \\ 1 - d_i.\epsilon, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

Here, d_i is the degree of the corresponding vertex, i.e., the number of its neighbors. As discussed in [23], the convergence of the algorithm is ensured in each non-bipartite non-regular graph when the mixing parameter ϵ is selected from the interval described in 8:

$$0 < \epsilon \leq \frac{1}{d_{max}} \tag{8}$$

Here, d_{max} is the degree of the best-connected vertex in a graph, so it can be represented using 9:

$$d_{max} = \max_i \{d_i\} \tag{9}$$

2.2 Analyzed Stopping Criteria

In this subsection, we introduce the implemented stopping criteria for bounding the execution of AC.

The first analyzed stopping criterion is proposed in [13] and poses a fully-distributed approach, i.e., no global information is necessary for its proper functioning. It is determined by two constants, namely *accuracy* and *counter threshold*, which are the same for each sensor node and preset before AC begins. Moreover, each sensor node has its own *counter*, which is initiated with "0" at each node. The principle of the stopping criterion is based on the calculation of the finite difference between the inner states at two consecutive iterations. If the finite difference is smaller than *accuracy*, the corresponding sensor node increments its *counter* by "1". If not, *counter* is reset[4] regardless of its current value. When the finite difference is smaller than *accuracy* more times[5] in a row, the algorithm is completed at the corresponding sensor node. Thus, this sensor node does not participate in AC and update its inner state any longer.

The other implemented stopping criterion from [6] is a centralized approach, i.e., it requires global information for its proper functioning. This stopping criterion is determined by only one constant *accuracy*, which is preset and the same for each sensor node again. Its principle lies in comparing the maximum and the

[2] Here, k is the label of an iteration, and $k = 0$ represents the initial inner state

[3] referred to as the mixing parameter ϵ

[4] i.e., set to "0"

[5] equal to the value of *counter threshold*

minimum from all the inner states with *accuracy*. The algorithm is globally stopped at the first iteration when the difference between the maximum and the minimum is smaller than the value of *accuracy*. Thus, the algorithm is executed until the condition described by 10 is met:

$$| max\{\mathbf{x}(k)\} - min\{\mathbf{x}(k)\} | < accuracy \tag{10}$$

3 Applied Methodology

In this section, we introduce the applied methodology and the used metrics for performance evaluation.

As mentioned earlier, AC described by the Perron Matrix meets all three convergence conditions in non-bipartite non-regular graphs when its mixing parameter ϵ takes a value from the interval (8). Thus, we select these four initial configurations for evaluation:

- $\epsilon = \{0.25 \cdot \frac{1}{d_{max}}, 0.5 \cdot \frac{1}{d_{max}}, 0.75 \cdot \frac{1}{d_{max}}, \frac{1}{d_{max}}\}$

These values are furthermore abbreviated as $\epsilon = \{$ 0.25, 0.5, 0.75, 1$\}$. As it is not likely that real-life WSNs are bipartite regular graphs [21], we omit these critical topologies from our analyzes.

As mentioned above, we bound the execution of AC by the stopping criteria from [6, 13]. In Section 2.1, it is stated that the fully-distributed stopping criterion from [13] is determined by two preset constants, whose values take these following values in our analyzes:[6]

- *accuracy* = $\{10^{-2}, 10^{-4}, 10^{-5}, 10^{-6}\}$
- *counter threshold* = {3, 5, 7, 10, 20, 40, 60, 80, 100}

Moreover, as mentioned earlier in this paper, we compare two scenarios related to this stopping criterion: either the initial inner states or the final estimates are multiplied by the network size n.

Furthermore, the performance achieved with the implemented stopping criterion from [13] is compared to MSE and the convergence rate when the centralized stopping criterion from [6] is implemented. As stated in the previous section, it is determined by *accuracy*, which takes the same values as the examined fully-distributed approach, i.e.:

- *accuracy* = $\{10^{-2}, 10^{-4}, 10^{-5}, 10^{-6}\}$

In this paper, AC is analyzed over RGGs of dense connectivity[7]. We generate 30 RGGs with unique topologies that are formed by 200 vertices each. A representative of the generated graphs is shown in Figure 2.

In our analyzes, each vertex initiates its inner states with a randomly generated scalar value of the standard Gaussian distribution, describe by 11:

$$x_i(0) \sim N(0, 1), \ for \ \forall v_i \in \mathbf{V} \tag{11}$$

To evaluate the performance of the algorithm, we apply two metrics for this purpose, namely the mean square error (MSE) and the convergence rate expressed as the iteration number necessary for the consensus achievement. In all the presented figures, we show MSE/the convergence rate averaged over 30 RGGs.

The first metric, MSE, is applied for precision evaluation of the final estimates and is defined in 12, [26]:

$$\text{MSE} = \frac{1}{n} . \sum_{i=1}^{n} \left(x_i(k_l) - \mathbf{1}^{\mathrm{T}} \times \frac{\mathbf{x}(0)}{n} \right)^2 \tag{12}$$

Here, k_l is the label of the iteration when the algorithm is completed at each vertex. The other metric, i.e., the number of the iterations, is applied in order to identify how long the algorithm has to be executed until the consensus upon the sum of all the initial inner states is achieved.

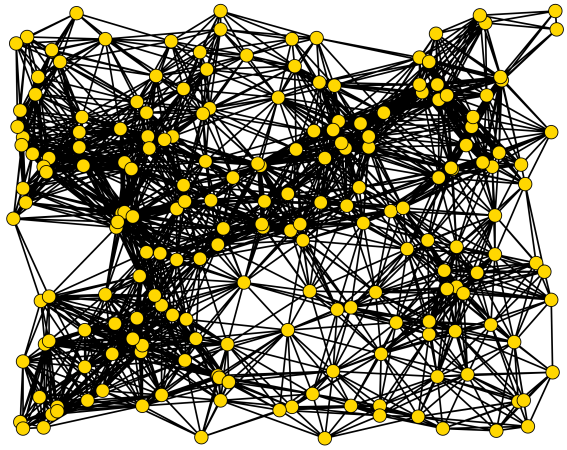


Figure 2: Representative of generated RGGs

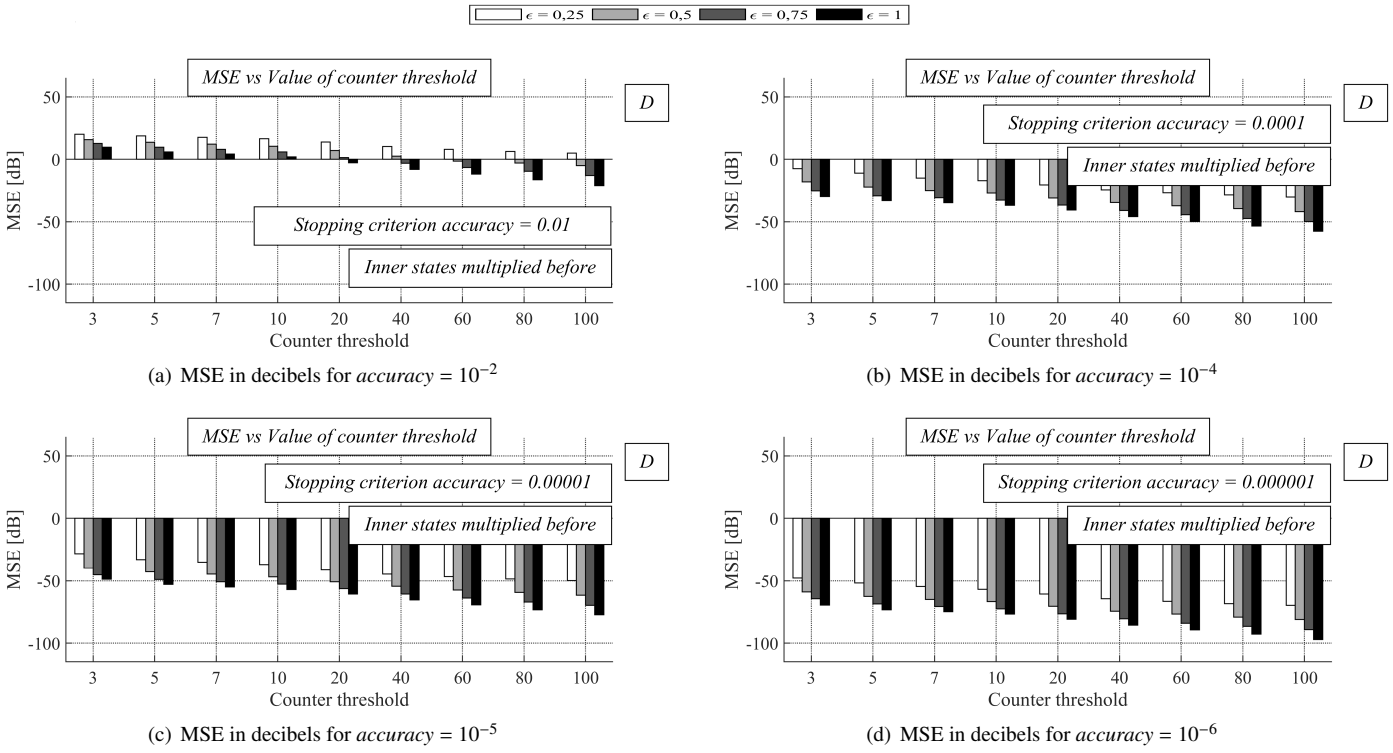
4 Experiments and Discussion

In this section, we present the results obtained in Matlab2016a and Matlab2018b and discuss the character of the depicted functions and observable phenomena.

In the first experiment, we analyze the precision of the final estimates quantified by MSE when AC is bounded by the stopping criterion presented in [13]. As already mentioned, we examine two scenarios: either the initial states (i.e., before AC begins – referred to as scenario 1) or the final estimates (i.e., after AC is completed – referred to as scenario 2) are multiplied by the network size n. From the results shown in Figure 3 and 4, we can observe in both examined scenarios that a decrease in the value of *accuracy* (see (a), (b), (c), and (d) in both figures for the results achieved for diffent values of *accuracy*) and an increase in the value of *counter threshold* ensure that MSE is smaller for each analyzed mixing parameter ϵ. Moreover, it can be seen that also an increase in the value of the mixing parameter ϵ results in a decrease in MSE for each *accuracy* and *counter threshold*. Furthermore, in scenario 2, we can see that

[6] the values 10^{-1}, 10^{-3} are omitted in order to ensure good readability of the paper

[7] in [1], also an analysis of AC for distributed summing in sparsely connected RGGs is provided, but it is not extended in this paper in order to achieve better readability of the paper



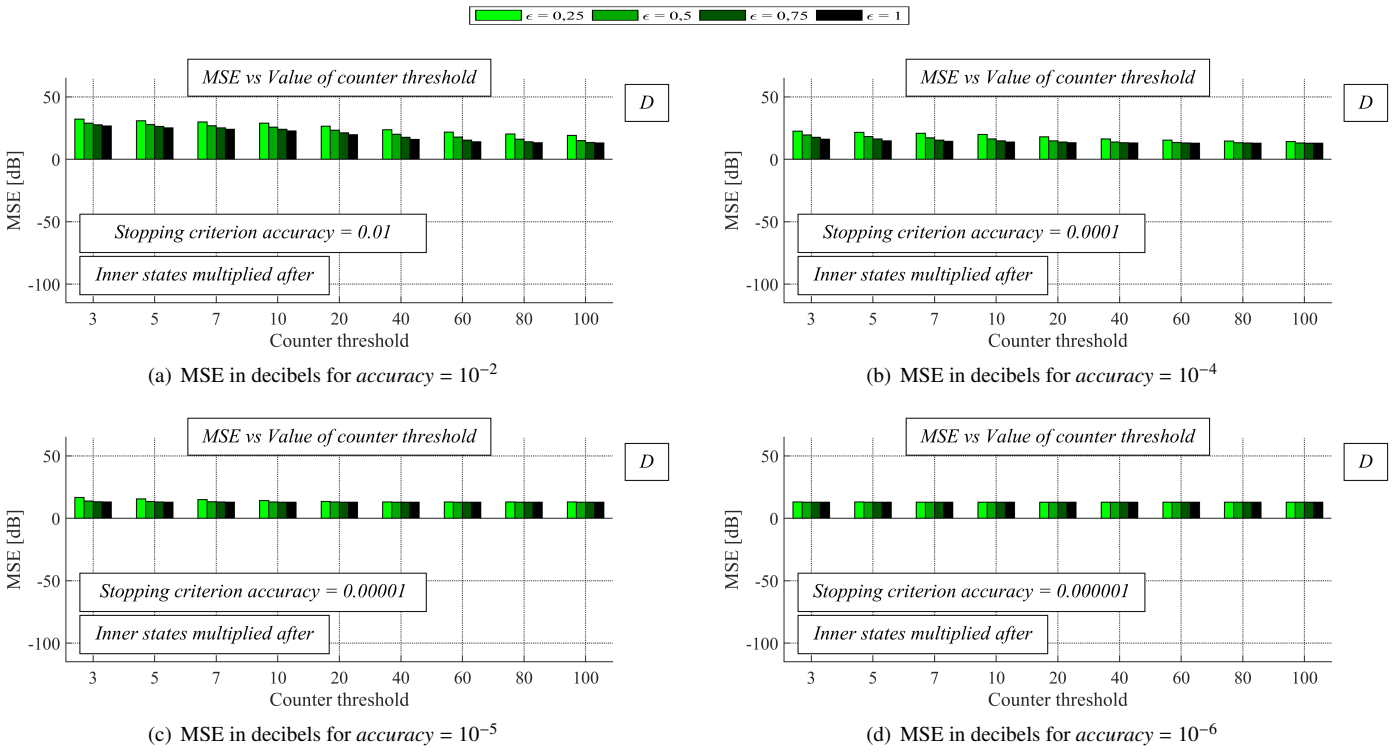
(a) MSE in decibels for *accuracy* = 10^{-2}

(b) MSE in decibels for *accuracy* = 10^{-4}

(c) MSE in decibels for *accuracy* = 10^{-5}

(d) MSE in decibels for *accuracy* = 10^{-6}

Figure 3: MSE [dB] averaged over 30 dense RGGs – fully-distributed approach is implemented, initial inner states are multiplied by network size



(a) MSE in decibels for *accuracy* = 10^{-2}

(b) MSE in decibels for *accuracy* = 10^{-4}

(c) MSE in decibels for *accuracy* = 10^{-5}

(d) MSE in decibels for *accuracy* = 10^{-6}

Figure 4: MSE [dB] averaged over 30 dense RGGs – fully-distributed approach is implemented, final estimates are multiplied by network size

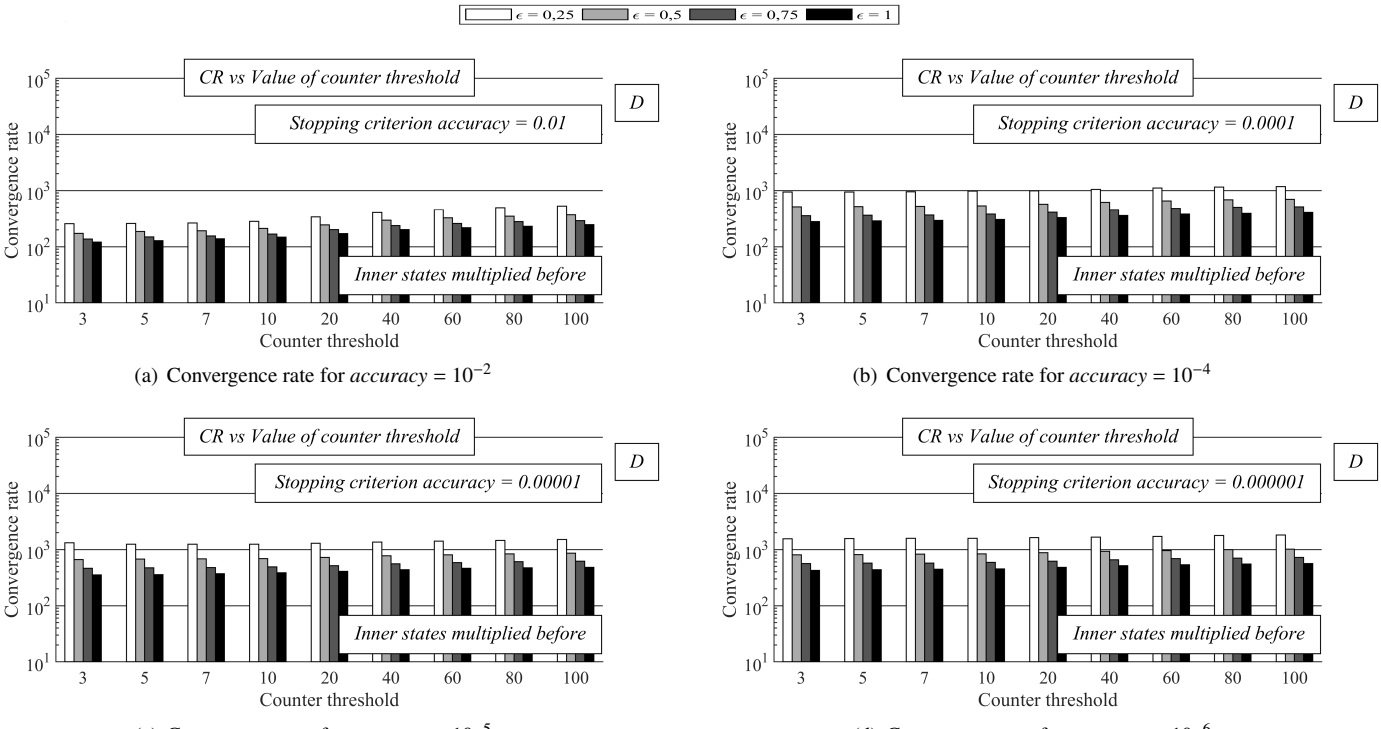
the precision of the final estimates is much smaller than in scenario 1, and a decrease in *accuracy* (see (a), (b), (c), and (d)) and an increase in *counter threshold* have a significantly smaller impact on an increase in the precision compared to scenario 1.

In the next experiment, we analyze the convergence rate expressed as the number of the iterations for the consensus achievement in both scenarios again. From the presented results in Figure 5 and 6, it can be seen that a decrease in the values of *accuracy* (see (a), (b), (c), and (d)) and an increase in the values of *counter threshold* result in a deceleration of the algorithm (i.e., AC requires more iterations to be completed) for each examined mixing parameter ϵ in both analyzed scenarios. Again, it is seen that AC with a larger mixing parameter ϵ achieves a higher performace like in the previous analyzes. However, unlike the previous analysis, AC in scenario 2 outperforms AC in scenario 1 for each value of both *accuracy* and *counter threshold*, and therefore, mutliplying the initial states with the network size n ensures higher precision of the final estimates, but a higher convergence rate is achieved in scenario 2.

In the last experiment, we test AC with the centralized stopping criterion from [6]. In Figure 7 and 8, we depict MSE and the convergence rate expressed as the number of the iteration for the consensus for same *accuracy* as in the previous two experiments. From Figure 7, it can be observed that a decrease in *accuracy* (see (a), (b), (c), and (d)) results in a decrease in MSE, and so, the precision of the final estimates is increased. Compared to the analysed fully-distributed approach, we can see that the value of the mixing parameter ϵ has only a marginal impact on MSE regardless of *accuracy* of the stopping criterion (in fact, increasing ϵ ensures a small increase in the performace). Regarding the convergence rate, a higher value of the mixing parameter ϵ ensures a higher convergence rate just like in the previous analyzes. Also, it is seen that a decrease in *accuracy* (see (a), (b), (c), and (d)) results in a lower convergence rate; therefore, more iterations are necessary for AC to be completed.

In the two following paragraphs, we turn our attention to a comparison of the centralized and the fully-distributed stopping criterion, which are analyzed in the previous parts - in all the comparisons, the results achieved for $\epsilon = 1$ are examined (i.e., the results achieved by the best performing initial configuration). From Figure 3, Figure 4, and Figure 7, where the precision of the final estimates is analyzed, we can see that a decrease in *accuracy* (see (a), (b), (c), and (d)) results in a decrease in MSE in all three cases. Also, it is seen that the value of the mixing parameter ϵ has only a marginal impact on MSE when the centralized stopping criterion is applied in contrast to the fully-distributed approach. In addition, in Figure 4, it is seen that the value of the mixing parameter ϵ less affects the precision of the final estimates for lower values of *accuracy* and higher values of *counter threshold*. Furthermore, in the first case, when AC is bounded by the fully-distributed stopping criterion in scenario 1, MSE is from the range <-97.12 dB – 9.63 dB>, meanwhile, in scenario 2, MSE takes the values from this interval <12.95 dB – 26.67 dB>. In the third case, when the centralized stopping criterion is implemented, the values of MSE are from <-131.80 dB – -52.07 dB>. Therefore, the centralized stopping criterion significantly outperforms the fully-distributed approach in terms of the estimation precision. Even though this approach achieves a significantly higher performance according to the MSE-metric, it is less appropriate for real-life implementations since it requires time-variant global information – the maximum and the minimum from all the current inner states at each iteration.

In terms of the convergence rate expressed as the number of



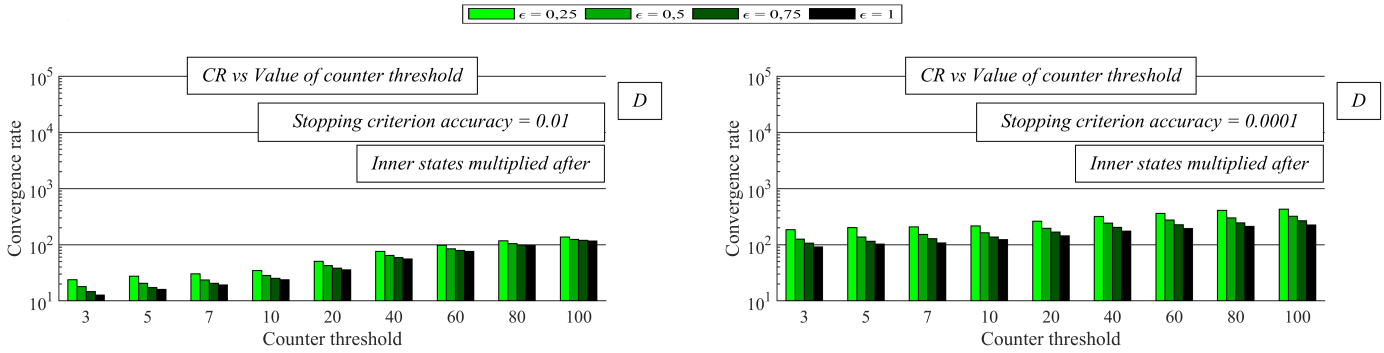
(a) Convergence rate for *accuracy* = 10^{-2}

(b) Convergence rate for *accuracy* = 10^{-4}

(c) Convergence rate for *accuracy* = 10^{-5}

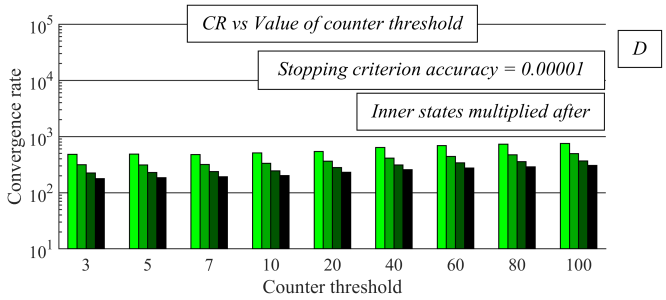
(d) Convergence rate for *accuracy* = 10^{-6}

Figure 5: Convergence rate averaged over 30 dense RGGs – fully-distributed approach is implemented, initial inner states are multiplied by network size

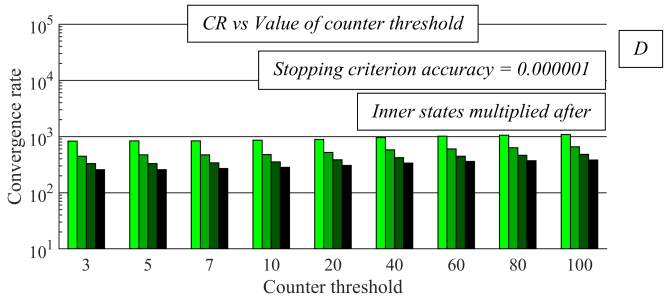


(a) Convergence rate for *accuracy* = 10^{-2}

(b) Convergence rate for *accuracy* = 10^{-4}

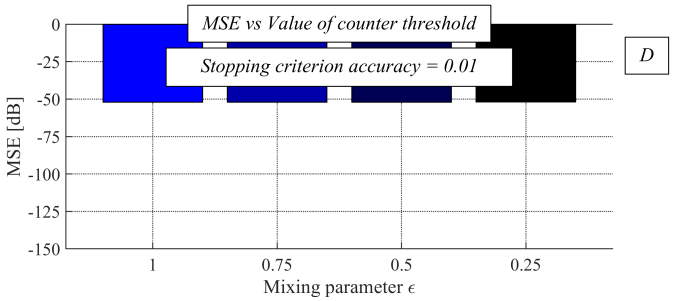


(c) Convergence rate for *accuracy* = 10^{-5}

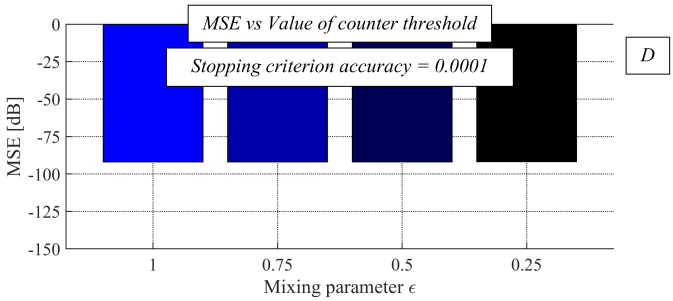


(d) Convergence rate for *accuracy* = 10^{-6}

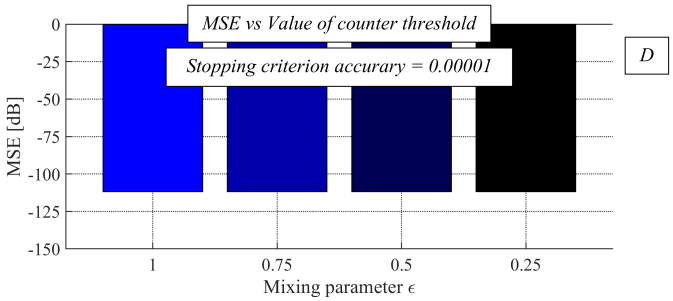
Figure 6: Convergence rate averaged over 30 dense RGGs – fully-distributed approach is implemented, final estimates are multiplied by network size



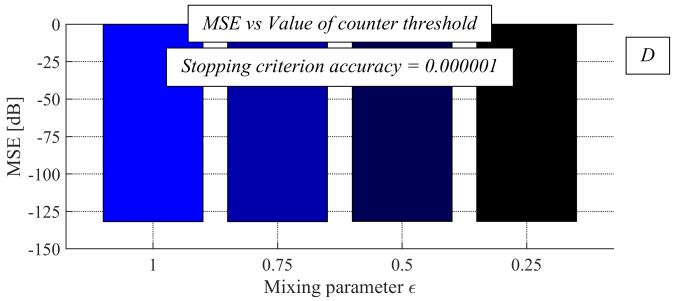
(a) MSE in decibels for *accuracy* = 10^{-2}



(b) MSE in decibels for *accuracy* = 10^{-4}

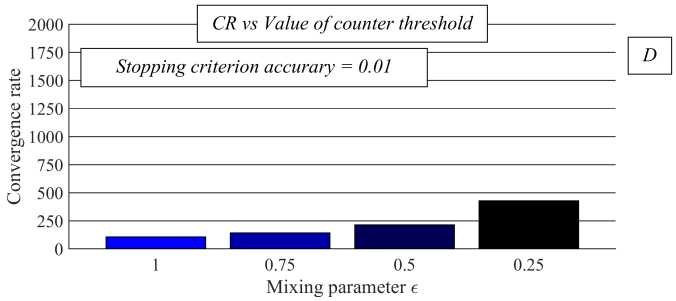


(c) MSE in decibels for *accuracy* = 10^{-5}

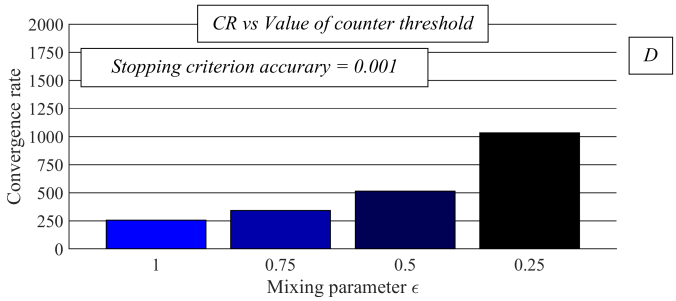


(d) MSE in decibels for *accuracy* = 10^{-6}

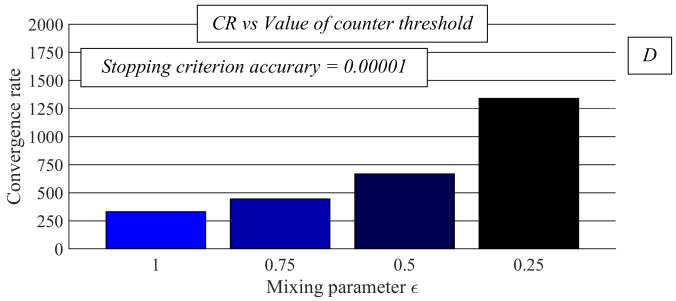
Figure 7: MSE [dB] averaged over 30 dense RGGs – centralized approach is implemented



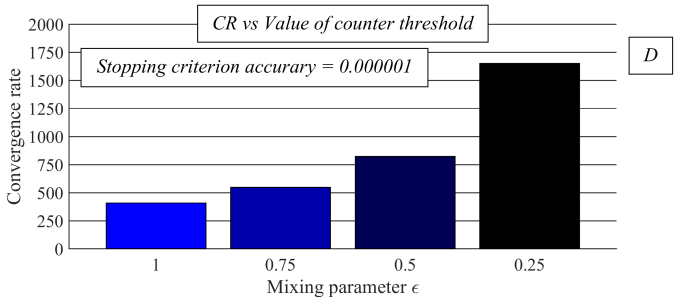
(a) Convergence rate for *accuracy* = 10^{-2}



(b) Convergence rate for *accuracy* = 10^{-4}



(c) Convergence rate for *accuracy* = 10^{-5}



(d) Convergence rate for *accuracy* = 10^{-6}

Figure 8: Convergence rate averaged over 30 dense RGGs – centralized approach is implemented

the iterations necessary for the consensus achievement (the results from Figure 5, Figure 6, and Figure 8 are compared), we can see that a decrease in *accuracy* (whereby a higher precision is achieved – see (a), (b), (c), and (d)) causes a deceleration of the algorithm in all three cases again. In the case, when AC is bounded by the fully-distributed stopping criterion in scenario 1, the convergence rate is from <121 - 560.4>. In scenario 2, the convergence is from the following interval <12.63 - 383.1>, and, in the third case, when the centralized stopping criterion is implemented, the convergence rate takes the following values <105.9 – 408.50>. So, it is seen that the highest convergence rate is achieved by the fully-distributed stopping criterion when the final estimates are multiplied by the network size n.

5 Conclusion

In this paper, we address AC for distributed summing whose execution is bounded by either a centralized or a fully-distributed stopping criterion. The results from our experiments show that decreasing *accuracy* ensures a higher precision of the final estimates but at a cost of a deceleration of the algorithm regardless of the applied stopping criterion. The precision of the final estimates can be improved also by increasing *counter threshold* when the fully-distributed stopping criterion is implemented. Moreover, it is proven that an increase in the mixing parameter ϵ optimizes the algorithm in terms of both the precision and the convergence when AC is bounded by the fully-distributed stopping criterion. In the case of the implementation of the centralized one, the value of ϵ has only a marginal impact on the precision of the final estimates, however, its higher values ensure a higher convergence rate like in the case of the fully-distributed stopping criterion. The fully-distributed stopping criterion achieves a higher precision when the initial inner states are multiplied by the network size n instead of the final estimates. However, on the other hand, multiplying the final estimates with n results in a higher convergence rate. Furthermore, it is proven that the centralized stopping criterion achieves higher performance in terms of the precision, however, this approach is less suitable for real-life implementations because it requires the information about the maximum and the minimum from all the current inner states at each iteration for its proper operating. The fastest is AC with the fully-distributed stopping criterion when the final estimates are multiplied by the network size n. On the other hand, in this case, the precision of the final estimates is significantly lower.

Acknowledgment

This work was supported by the VEGA agency under the contract No. 2/0155/19 and by COST: Improving Applicability of Nature-Inspired Optimisation by Joining Theory and Practice (ImAppNIO) CA 15140. Since 2019, Martin Kenyeres has been a holder of the Stefan Schwarz Supporting Fund.

References

[1] M. Kenyeres, J. Kenyeres, I. Budinska, "Distributed linear summing in wireless sensor networks" in IEEE 17th World Symposium on Applied Machine Intelligence and Informatics (SAMI 2019), Herlany, Slovakia, 2019. https://doi.org/10.1109/SAMI.2019.8782782

[2] H. Yetgin, K. T. K. Cheung, M. El-Hajjar, L. Hanzo, "A survey of network lifetime maximization techniques in wireless sensor networkss" IEEE Commun. Surv. Tutor., **19**(2), 828–854, 2017. https://doi.org/10.1109/COMST.2017.2650979

[3] B. Rashid, M. H. Rehnami, "Applications of wireless sensor networks for urban areas: a survey" J. Netw. Comput. Appl., **60**, 192–219, 2016. https://doi.org/10.1016/j.jnca.2015.09.008

[4] T. Qiu, A. Zhao, F. Xia, W. Si, D. O. Wu, "ROSE: robustness strategy for scale-free wireless sensor networks" IEEE/ACM Trans. Netw., **25**(5), 2944–2959, 2017. https://doi.org/10.1109/TNET.2017.2713530

[5] B. Krishnamachari, D. Estrin, S. B. Wicker, "The impact of data aggregation in wireless sensor networks" in 22nd International Conference on Distributed Computing Systems Workshops (ICDCSW 2002), Vienna, Austria, 2002. https://doi.org/10.1109/ICDCSW.2002.1030829

[6] M. Kenyeres, J. Kenyeres, V. Skorpil, "The distributed convergence classifier using the finite difference" Radioengineering, **25**(1), 148–155, 2016. https://doi.org/10.13164/re.2016.0148

[7] Y. Liu, M. Dong, K. Ota, A. Liu, "ActiveTrust: secure and trustable routing in wireless sensor networks" IEEE Trans. Inf. Forensics Secur., **11**(9), 2013–2027, 2016. https://doi.org/10.1109/TIFS.2016.2570740

[8] D. Izadi, J. H. Abawajy, S. Ghanavati, T. Herawan, "A data fusion method in wireless sensor networks," Sensors, **15**(2), 2964–2979, 2015. https://doi.org/10.3390/s150202964

[9] L. Xiao, S. Boyd, S. Lall, "A scheme for robust distributed sensor fusion based on average consensus" in 4th International Symposium on Information Processing in Sensor Networks (IPSN 2005), Los Angeles, CA, USA, 2005.

[10] G. Stamatescu, I. Stamatescu, D. Popescu, "Consensus-based data aggregation for wireless sensor networks" Control. Eng. App. Inf., **19**(2), 43–50, 2017.

[11] A. K. Idrees, W. L. Al-Yassen, M. A. Taam, O. Zahwe, "Distributed data aggregation based modified K-means technique for energy conservation in periodic wireless sensor networks" in 2018 IEEE Middle East and North Africa Communications Conference (MENACOMM 2018), Jounieh, Lebanon, 2018. https://doi.org/10.1109/MENACOMM.2018.8371007

[12] Q. Chen, H. Gao, Z. Cai, L. Cheng, J. Li, "Distributed low-latency data aggregation for duty-cycle wireless sensor networks" IEEE/ACM Trans. Netw., **26**(5), 2347–2360, 2018. https://doi.org/10.1109/TNET.2018.2868943

[13] M. Kenyeres, J. Kenyeres, M. Rupp, P. Farkas, "WSN implementation of the average consensus algorithm" in 17th European Wireless Conference 2011 (EW 2011), Vienna, Austria, 2011.

[14] D. Merezeanu, M. Nicolae, "Consensus control of discrete-time multiagent systems," U. Politeh. Buch. Ser. A, **79**(1), 167–174, 2017. https://doi.org/10.1109/TNET.2018.2868943

[15] P. Jesus, C. Baquero, P. S. Almeida, "A survey of distributed data aggregation algorithms" IEEE Commun. Surv. Tutor., **17**(1), 381–404, 2015. https://doi.org/10.1109/COMST.2014.2354398

[16] A. Olshevsky, "Linear time average consensus and distributed optimization on fixed graphs" SIAM J. Control. Optim., **55**(6), 3990–4014, 2017. https://doi.org/10.1137/16M1076629

[17] G. Oliva, R. Setola, C. N. Hadjicostis, "Distributed finite-time average-consensus with limited computational and storage capability" IEEE Trans. Control. Netw. Syst., **4**(2), 380–391, 2017. https://doi.org/10.1137/16M1076629

[18] L. Faramondi, R. Setola, G. Oliva, "Performance and robustness of discrete and finite time average consensus algorithms" Int. J. Syst. Sci., **49**(12), 2704–2724, 2018. https://doi.org/10.1080/00207721.2018.1510059

[19] J. Gutierrez-Gutierrez, M. Zarraga-Rodriguez, X. Insausti, "Analysis of known linear distributed average consensus algorithms on cycles and paths" Sensors, **18**(4), 1–25, 2018. https://doi.org/10.3390/s18040968

[20] T. C. Aysal, B. N. Oreshkin, M. J. Coates, "Accelerated distributed average consensus via localized node state prediction" IEEE Trans. Signal Process., **57**(4), 1563–1576, 2009. https://doi.org/10.1109/TSP.2008.2010376

[21] V. Schwarz, G. Hannak, G. Matz, "On the convergence of average consensus with generalized Metropolis-Hasting weights" in 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2014), Florence, Italy, 2014. https://doi.org/10.1109/ICASSP.2014.6854643

[22] L. Xiao, S. Boyd, "Fast linear iterations for distributed averaging" Syst. Control. Lett., **53**(1), 65–78, 2004. https://doi.org/10.1016/j.sysconle.2004.02.022

[23] N. E. Manitara, C. N. Hadjicostis, "Distributed stopping for average consensus in directed graphs via a randomized event-triggered strategy" in 6th International Symposium on Communications, Control and Signal Processing (ISCCSP 2014), Athens, Greece, 2014. https://doi.org/10.1109/ISCCSP.2014.6877918

[24] S. V. Macua, C. M. Leon, J. S. Romero, S. S. Pereira, J. Zazo, A. Pages-Zamora, R. Lopez-Valcarce, S. Zazo, "How to implement doubly-stochastic matrices for consensus-based distributed algorithms" in 2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM 2014), A Coruna, Spain, 2014. https://doi.org/10.1109/SAM.2014.6882409

[25] V. Schwarz, G. Matz, "Nonlinear average consensus based on weight morphing" in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), Kyoto, Japan, 2012. https://doi.org/10.1109/ICASSP.2012.6288578

[26] S. S. Pereira, A. Pages-Zamora, "Mean square convergence of consensus algorithms in random WSNs" IEEE Trans. Signal Process., **58**(5), 2866–2874, 2010. https://doi.org/10.1109/TSP.2010.2043140

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 28-33 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Centralized System of Universities Learning Materials

Ruslan Vynokurov*, Volodymyr Tigariev, Oleksii Lopakov, Kateryna Kirkopulo, Olena Pavlyshko

Odessa National Polytechnic University, Institute of Industrial Technologies Design and Management, 65044, Ukraine

ARTICLE INFO

Article history:
Received: 15 January, 2020
Accepted: 11 April, 2020
Online: 03 May, 2020

Keywords:
Distance Learning
Education
Studying
Website
VR
AR

ABSTRACT

This article considers the creation of a program / website as an example of a centralized system for all universities in order to more easily familiarize an applicant and / or student with the internal politics and capabilities of the university. Existing problems in the training system are examined, why distance learning is relevant, what problems exist in it and how they can be solved. Based on the proposed distance learning projects, an argument is built on its meaninglessness without further advancement, and how a website concept can help solve this problem. The concept of the website and its possible implementation using existing analogues in other areas that successfully complete the tasks are considered. Theoretically, using the proposed methods, the concept will simplify the choice of a further place of study for schoolchildren and applicants. The method consists in parsing the desired university website into the necessary categories: a number of specialties, disciplines, foreign programs, distance learning materials and so on. This will allow you to create a library of universities, their materials and open data (specialties, disciplines), without coming to excesses, such as creating a video hosting service for broadcasting recorded disciplines, less data to be stored, and more.

1. Introduction

This paper is an extension of work originally presented in IEEE PICS&T 2018 [1].

The purpose of this article is to consider the possibility of creating a website that will combine higher education institutions with the goal of easier navigation in higher education.

The original article examined why the decision to use the Unreal Engine 4 (https://www.unrealengine.com/en-US/) game engine is relevant to create a platform. Thanks to blueprint technology, which is a system of nodes with built-in code scripts, you can easily perform many tasks without knowing the syntax of a programming language, in this case C++. Saving time for studying the syntax and focusing forces on the program logic, it becomes possible to devote more time directly to the program itself, which positively affects the quality.

Also in the original article, the creation of a workflow for filling the platform with materials was considered: what components are needed to create a working system, what programs and their bounds were proposed for compiling the content of the proposed lessons, a conceptual example used at the Odessa National Polytechnic University was considered. All this was done in order to describe one of the possible options for creating such a platform/hub, where students could find the materials they need and navigate the relevance of knowledge that the university has been presenting during the studying.

This article focuses on the consideration of the previously described problem, its expansion and argumentation. An approach was taken to collect information from third-party sources and combine them to provide more specific structured information to applicants through the website. The process of the platform creation was partially considered [1,2].

At the moment, students are often not aware of what universities are like. According to the experience of the authors who live in Ukraine, at the start of student enrollment, different universities arrange open house days when you can go in and look at the place where the applicant thinks to continue his studies. According to this event, the student receives a plot of information about who he will study with, who will teach him and what he will be able to achieve. Often, since in Ukrainian universities a shortage of students in certain specialties is a frequent occurrence, it is customary to embellish a university story and indicate only the advantages of the specialty and/or specialization, without informing applicants about possible shortcomings. The

*Ruslan Vynokurov, Email: ruslan.vinokurov14@gmail.com

disadvantages can be in different areas: poor teaching staff, the relevance of the teaching materials presented, poor learning conditions and more.

There is also the problem of choice, which should solve the confines between graduation and the beginning of a university.

The problems that prevent students from choosing their future profession will be considered below.

2. Main Problem

There are a number of problems that do not allow obtaining the necessary knowledge for a person. It is also necessary to indicate that these problems are considered in several areas: preparation of the applicant for admission to the university and student gaining knowledge in the university environment. According to the authors, there are two most common problems:

- The problem of awareness.
- The problem of the existing education system.

2.1. Awareness problem

The authors' experience is based on a model of the Bologna system, as well as on the basis of statistics from some universities in the Russian Federation. The moment of transition from school to a higher educational institution often raises many questions, since this stage is constantly updated and supplemented by various acts. All these updates are briefly explained, but the busyness of schoolchildren with the upcoming exams does not make it possible to immerse themselves in the question completely.

The exam procedure itself is relatively simple, but preparing for them can entail enormous costs in time, which limits students in reflecting on who they want to become in the future. Short time for studying the material and pressure from the family and the teaching staff leads to procrastination among schoolchildren, which affects their general level of anxiety, productivity and academic performance [3]. Such employment and a psychologically difficult situation do not allow the student to adequately assess their strengths and weaknesses, which would allow to come to the correct decision regarding the choice of specialty and specialization.

According to the considered statistics, only 11% of schoolchildren in grades 8-9 are determined with the place of forthcoming education. 54.1% is determined only by the end of grade 11. 5.2% of students initially knew where they would go. 19.7% are determined either immediately before admission, or do not do it at all. The choice is complicated by a large number of criteria and a variety of conditions, thanks to through which the choice of university is made: parents' advice, teachers' advice, print university advertising, television and radio advertising, open day at the university, conversations with university students, speeches, meetings, conversations with teachers, job fairs, university rankings in the media and more [4]. Due to the fact that applicants cannot decide on the choice of a profession at school, they are trying to find the specialty that is relevant and necessary at the time of selection — the trend.

Often, professions are selected on the basis of superficial knowledge about this type of activity. Each specialty has its own requirements for a person in different fields: level of knowledge, stress tolerance, responsibility, level of emotional intelligence, sociability, teamwork and other human qualities. This can be called a characteristic of a single profession, which for the most part should be observed in candidates. Many people make an unconscious choice of their future profession, since the characteristic is a huge amount of details and, often, the only way to understand whether you can relate to this specialty is an internship, which can be completed with an average level of knowledge in the specialty, which suggests studying in this area. Therefore, those 54.1%, as well as 19.7% of schoolchildren who make decisions at the last moment, cannot have a complete picture of what they need to study in order to become a specialist in a certain kind of activity, since this requires a lot of amount of time. This leads to poor-quality training of students due to the lack of their motivation to study further. And this, in turn, to work not in the specialty and lack of qualifications in another kind of activity.

2.2. Education system problem

Countries should try to develop in the field of education as fast as countries with similar capabilities do [5]. The main problem in the development of distance learning is an incorrect education system. According to statistics from the American Federation's National Policy Summit for Children, online schools in their early days did not achieve proper exams due to funding for a non-working education system. From the very beginning, virtual schools were suitable only for some disciplines [6], but with the advent of new mechanics it became stronger in others, which gradually leads to the popularity of this approach for learning.

In the CIS countries, this situation lies in the lag of the teaching methodology, which is manifested in the students' lack of interest, since the training system is not clear and complex, which leads to a lack of understanding of the need for a particular discipline and/or specialty as a whole. This is due to the formalism of the materials presented. It is expressed in a mechanical, passive, insufficiently meaningful assimilation by students of teaching material, in memorizing verbal formulas by them, devoid of specific content, in the inability to connect the knowledge gained with life [7]. This educational system was used in the USSR and developed to the maximum possible to meet the requirements of the ruling party. Subsequently, the education system was not changed and only updated the knowledge provided, which led to the depletion of the system and its inability to improve further, namely, to include modern technologies as part of the learning process, such as AR/VR/MR and so on [8].

3. Information About Distance Learning

At the moment, distance learning is becoming more and more popular in the world, which helps many students to gain knowledge without being in an educational institution. Distance learning is gradually becoming the standard throughout the world. This can be seen by the number of queries in popular search engines. For example, according to the data from the multi-functional SEO platform Serpstat, on google.com the query "distance learning" is entered into the search line 133900 times a month (on average for 2019), and the phrase "online school" is 255300 times under the same conditions. Now the term "online school" is becoming popular, which consists in distance learning through various communication platforms and video conferencing

of the teacher. Special business accelerators are being created, which are aimed at helping start-ups and companies related to this type of studying and helping to transform a successful business out of it through templates and promotion schemes.

The largest universities in the world have long possessed this type of education, for example: Boston, Florida and Arizona State Universities, the University of Wisconsin at Madison, as well as the State University of Pennsylvania in the USA. With the help of this system, the University of Florida allowed 400000 students to get education in 135 countries of the world. There are examples in the UK: University of Liverpool, Suffolk University Campus, Ruskin University England, School of Oriental and African Studies, University of Manchester. The latter currently provides knowledge to about 40000 students only through distance learning in 154 countries around the world [5].

3.1. What is wrong with distance learning organization

Despite such local and international success, it is not possible to refer to the current distance learning - there is no primary source that would immediately tell about many available options from different universities and at the same time give access and description to it.

The above examples are well-known mastodons of science, which in any case will be found by applicants if they wish, because you can hear about them at school age as the most desirable places to study, but applicants absolutely do not know anything about other places, which also may be good for learning. Examples show that distance learning is valid, but it is impossible to confine oneself only to these mastodons - on the contrary, it is necessary to expand knowledge, it is necessary to make it clear that the world of education is much wider and more spacious, and does not close to the top ten most popular universities.

For example, the YouTube website (https://www.youtube.com/) is a great place to store distance learning videos, but it was not originally created for this. You can restrict access to video, as well as collect video playlists, but adding a description to each playlist to make it convenient is impossible due to lack of functionality. And it would be possible to create an analogue of Instant View (hereinafter referred to as the IV) from Telegram (https://telegram.org/) and still store the video on a service from Google, but at the same time pre-set and display it with the necessary data on the website.

The point is that it is necessary to fundamentally change the existing system and introduce a core that can connect all that is different that exists at the moment. Each person should have the opportunity to look at different universities and decide which one will be most suitable for him. This implies the lack of centralization of the final result, in this case, the showcase of distance learning and the university as a whole.

3.2. Why is this a problem

At this stage, an applicant who has the opportunity to enter any university should find universities on the Internet and check the availability of specialties, specializations, what he will be taught, who he will become in the end, what perspectives the industry has and much more. It takes time to surf on the Internet, which students do not have, as it was described earlier. The problem is that the more responsible the student approaches the question of exams, the more irresponsible he is to his future. It is quite difficult to find a balance, since both questions are complex and difficult to study, so the idea of centralizing information about universities from one point seems quite reasonable.

3.3. How a solution to this problem affects on progress

When a website appears with an extensive list of universities, it becomes possible to compare their statistics: popular specialties, the relevance of discipline programs, teacher qualifications, etc.

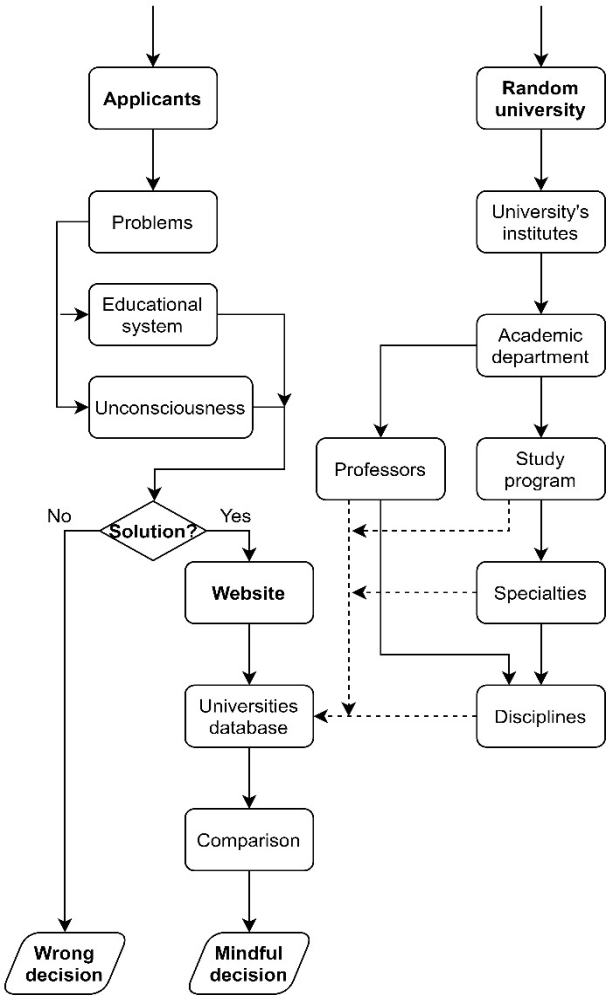


Figure 1: Scheme of each applicant choice and how the solution should be implemented.

Thanks to this website, students can easily compare and contrast universities: see different and general, study programs, with which foreign partners there is cooperation, foreign programs, opportunities for graduation, with which companies the educational institution cooperates. Also, applicants gain the opportunity to see other universities, the existence of which they might not know. Since the website will also have distance learning materials, this allows applicants to review the teaching methodology, compare with other universities and make an informed decision about admission to a particular educational institution.

This implies a great knowledge of the choice of a university, as well as a dry balance of the necessary information without frills,

which allows a sober assessment of the proposed educational institution. As a result, students study in more relevant specialties for them at the universities most suitable for this, which is what the research of this article seeks. The overall idea is described on the Figure 1.

4. Solution

The examples of solutions that allow you to learn much more efficiently and faster are shown below, but due to the lack of information about such solutions, few people think about their existence. At the same time, they can provide an incentive to create competitive software, or make additions to existing code, which will improve the UX and functionality of existing programs. This approach is widespread in the IT field and this is evidenced by such communities as GitHub, Stack overflow, the open source Unreal Engine 4, which can be supplemented by various plug-ins and the Blender 3D modeling program, which also has open source code and can be supplemented by functionality that users deem necessary.

All this speaks of the popularity of the approach when people, specialists in their field, are aware of this. That is why it is necessary to show and disseminate new techniques and show them to a wide audience, in which the proposed website concept can help.

4.1. Solution examples

There are different AR applications and many of them have very different goals. AR technology allows you to reproduce objects that are difficult to read in 3D models, which makes it easier to accept abstract and complex content. This is useful for those who remember visually and who need a translation of the theory into a tangible concept. For example, Polytechnic University of Leiria, in Portugal introduced AR to math classes and students speak of it as a very useful, easy and interesting solution.

Nvidia Holodeck may be an example of VR technology, whose technology allows people from different parts of the planet to work together in virtual space, makes it possible to develop design, assembly lines, technological processes and much more, having their 3D counterparts in a virtual environment [1].

A team from Stanford University created a project called STRIVR. It began as a VR training tool for their university football team. Subsequently, they found out about the project, and it turned into a platform for staff training and works with companies such as Walmart. United Rentals (hereinafter referred to as UR), which is the largest equipment rental company in North America, trains third-party sales representatives (hereinafter referred to as TPSR) using STRIVR technology. Using photos or videos in the classroom does not allow students to gain practical experience, because UR used the STRIVR immersion platform to create the next generation TPSR tutorial that uses VR technology. It includes 5 practical steps in VR, so that training takes place with a sense of presence directly at the workplace, but actually not there. Thus, they achieved an increase in efficiency and reduced training time by 40% [1].

4.2. Conclusions from the proposed concepts

Above are local, point solutions that allow you to gain an increase in the effectiveness of training using practical experience. This decision is relevant, as the current education system in the CIS has reached its maximum. At the moment, it cannot fully compare the modern requirements of social expectations and real educational results [8], which leads to a lag and less interest of students to study.

Also, these examples are relevant, since AR/VR/MR technologies are considered one of the most promising technologies of the 21st century, and also create new ways of presenting information and studying.

The disadvantage of this system is the lack of cooperation between such solutions, which does not allow to switch from one technology to another within the framework of one application. Having created a common basis that opens up access to any of these solutions, you can think about implementing a technology that will demonstrate different disciplines (where applicable), allowing students to understand what a single discipline is in a particular specialty and what it is manifests itself.

This technology allows to solve the problem of awareness in view of a more affordable solution for students and the opportunity to get some practical experience immediately before studying in this area.

Also, this system involves either a complete change in training programs, or their adaptation to a new approach, which solves the problem of the education system and allows to increase the effectiveness of training, a more relevant choice by students of future specialties and areas in which they will study.

4.3. Suggestion of solution

It is proposed to create a website. It will serve as the basis for comparing universities, searching for information about educational institutions, study programs in them. It is proposed to minimize the work and use what already exists, namely, materials and approaches.

Telegram messenger has IV technology, which was created specifically to display the content of other websites inside the messenger. When a user shares a website link inside Telegram, a bot instantly fires, which checks for the presence of a template for the website. A template is a set of rules that interprets a website as necessary, removing all unnecessary and leaving a purely necessary context in the form of text, links, tables, images. If such a template is found, the website opens inside the messenger without loading the page - instantly. It is noteworthy that any user can create a template and send it to the messenger developers for verification and subsequent approval. After confirming the suitability of the template, any page of the website will open according to the rules established in the template.

The idea of the proposed website is a similar approach. Instead of re-creating the content, you can duplicate it and interpret it in the right way, managing the content as you need.

This solution allows you not to create your own video hosting services, like YouTube, for storing university video data. Instead, you can simply link to existing pages and present information from the websites of universities and their individual projects, whether it's distance learning, or international programs, in one light and together in a user-friendly way.

This will allow you to view from one place all those existing lectures and materials, regardless of what they will be: an article, video, or the AR / VR / MR application. In the case of the latter, most likely, it will be necessary to install additional software recommended for their reproduction, but such a question is not included in the scope of this article.

The uniqueness and unification of the website is a guarantee of freedom in the implementation of projects (websites) of other universities. The existence of such a website will not limit the creators of university websites to change anything in the design or development approach.

5. Implementation

As mentioned earlier, the concept is a website that uses a similar IV technology created by Telegram.

The proposed website concept can solve the problems indicated in section 2. Since the concept collects a database of different places for higher education, having the ability to expand and modify the list of institutions and their characteristics according to the method below, the problem of awareness can be overcome with a single condition: the provision of such a basis to applicants in advance and with an explanation.

The system works like an online store with a large number of filters - what interests the applicant has selected: direction, specialty (final profession), country, tuition fees, other filters that will allow you to get a selection of final results. Provides a description of the specialty, possible previews of training materials for better reference. Applicants have a chance to look at a specialty and form an opinion about it, which makes it easier to make a decision.

The final result of solving the problem of awareness remains behind the desire of applicants to understand this issue, but the presented concept will simplify the task for them by collecting everything they need in a single database.

Initially, existing university materials will be presented, both on distance learning and lecture materials from various platforms, for example, YouTube. The website will initially focus on downloading various additional materials from external sources. This is due to the lack of servers and capacities that are required to service such networks. The website will only target educational materials from trusted sources that are either presented at the university, or distance learning materials. This is also the main difference from YouTube, Vimeo (https://vimeo.com/) and other similar services, where in addition to educational content, you can also find entertaining, which can distract the user. Thus, having the opportunity to accept materials from a variety of resources, the website acts as a place where only training materials will be collected.

5.1. Functionality and purpose of the site

To be able to open articles from a single resource, a template is created that transforms the data into the desired form and provides it in an IV which is a simplified form without such additional aspects as a list with other articles, website menu, advertising, and so on.

The transformation of the page of any resource is carried out in the following 4 steps:

- Determining which parts of a particular resource will be better suited to the conditions for creating the template;
- Parses the basis of the website for the main tags of HTML code to present them in accordance with the requirements of the template. In case an information element is found that does not meet the rules of the template, it is necessary to use the functions provided by Telegram that transform them into a suitable element;
- Remove all unnecessary content and provide a clean IV page for viewing using a dedicated feature;
- Check the template using 5-10 links to other articles of the same resource to make sure that all applied functions work correctly and there is no need to use others to display additional content (for example, replace one HTML tag with another, insert a paragraph tag inside text structures and so on).

According to a case study that was provided by Telegram based on an analysis of an article from medium.com, you can understand and take as a basic principle how the website content is transformed and its new look is compiled.

This approach is quite flexible and scalable. Also, the Telegram code is open and exposed on the GitHub resource, which allows you to use the acquired technologies as needed.

6. Related Work

At the moment, work is underway on the development of application authentication and the functionality of the hub start page, the second component of the proposed solution, which was presented in the original article [1]. To create authentication, a set of libraries is used to implement standard scenarios (hereinafter referred to as the SDK) of GameSparks (https://bit.ly/3besv6D), which allows us to develop user authorization logic in detail. The start page at the moment is a block that contains video, the main control panel and news.

For authentication, blueprint technology was used, which allowed us to create user registration and authentication. Figure 2 shows the blueprint used.

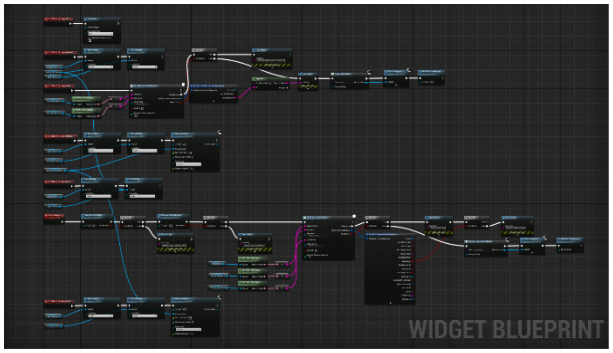


Figure 2: Authentication blueprint.

At this stage, the website has not yet been developed, since the final goal is constantly updated and optimized.

7. Future Work

As the upcoming works, the authors of this article see the development of the hub application, namely the creation of an application for playing their own AR/VR/MR projects and interaction with them. It will be considered creating a user account and saving information about it inside the application, the ability to edit it. Creating basic templates for learning materials, for example: video lecture, video with auxiliary files, AR project, VR project and more, in which all the necessary variables and components that create this template will be explained and recorded. Thanks to it, content of the same type will be able to collect in arrays and automatically load.

Viewing videos from YouTube, Vimeo and similar websites will be implemented, when creating the subsequent pages of the application.

Creation of the website described in this article where the functionality for comparing universities will be created. A bundle of a website and an application for synchronizing databases, since, for example, the GameSparks SDK stores all accounts inside its server.

The application will be created using Unreal Engine 4, which will fully realize the main goal of the article: quick familiarization of applicants with the specialties and professions taught by universities in order to make a more informed and rational choice of further education.

References

[1] V. Tigariev, R. Vinokurov, A. Pavlyshko, I. Sinko, "Informational Training System Mindgate" in 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), IEEE, 2019. https://ieeexplore.ieee.org/document/8632104

[2] V. Tigariev, R. Vynokurov, "Creation Of a Learning System Based On The Game Engine Unreal Engine 4 (article in Russian with an abstract in English)" in 2018 ICT 7th International Scientific and Technical Conference, Odessa, Ukraine, 2018. http://immm.opu.ua/files/archive/n4_v8_2018/immm_n4_v8_2018.pdf

[3] J.R. Ferrari “Self-handicapping by procrastinators: Effects of task importance and performance privacy” in 1990 Doctoral dissertation, Adelphi University, Garden City, NY, USA, 1990.

[4] G. Evdonin, S. Mamedova, “Modeling of choosing of university for entrants (article in Russian with an abstract in English)” in 2014 Управленческое консультирование, 2014. https://cyberleninka.ru/article/n/modelirovanie-protsessa-vybora-vuza-abiturientom-pri-postuplenii

[5] R. Vynokurov, "Расстояние Как Основная Проблема Информационного Обеспечения Образовательных Процессов" in 2019 TIRMIA/TI 5th International Scientific and Practical Conference of students, postgraduates and junior researches, Odessa, Ukraine, 2019.

[6] A. Molnar, G. Miron, C. Gulosino, C. Shank, C. Davidson, M. K. Barbour, L. Huerta, S. R. Shafer, J. K. Rice, D. Nitkin, “Virtual Schools in the U.S. 2017” in 2017 Boulder, CO: National Education Policy Center, 2017. https://nepc.colorado.edu/publication/virtual-schools-annual-2017

[7] O. Lebedev, “The End of Compulsory Education? (article in Russian with an abstract in English)” in 2016 Вопросы образования, 2016. https://cyberleninka.ru/article/n/konets-sistemy-obyazatelnogo-obrazovaniya

[8] K. Gusev, “XXIV съезд КПСС и отечественная историография” in 1971 Издательство Академии наук СССР, 1971. https://bit.ly/2uCNSxQ

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 28-34 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Ontological and Epistemological Issues of Studying the Development of Uzbekistan in the Context of Globalization from the Point of View of the Paradigm of Civilizations

Djurayeva Nigora Avazovna*

National University of Uzbekistan named after Mirzo Ulugbek, Tashkent Uzbekistan Doctoral student of the faculty of Social sciences, the Department of Philosophy and Logic, 702506, Uzbekistan

ARTICLE INFO

Article history:
Received: 19 August, 2019
Accepted: 05 February, 2020
Online: 09 March, 2020

Keywords:
Uzbekistan
Paradigm of civilizations
Ontology
The factors of development of the paradigm of civilizations
Strategy of development
Culture
Synergetic Approach
Geographical Factor
Space
Time
Value Factor
Modernization

ABSTRACT

In this scientific work, the basic synergetic principles of the development of the paradigm of civilizations are substantiated, an assessment of the time boundaries within which civilizations exist, the features of their own path of development of Uzbek culture are revealed. It is especially drawn attention to the sociocultural codes of civilization, a combination of which identifies the essence of Uzbek civilization. However, gives the reasons for the threat of the existence of civilizations and the prospects for their further development. The problem of the paradigm of civilizations interested many inquisitive minds of mankind, philosophical schools and trends. Its origins can be found in antiquity and in the Middle Ages. Dosocrats, Plato, Aristotle, Stoics, Epicureans, Neoplatonists, Thomas Aquinas, al-Farabi, Ibn Sina developed problems of society and history within the framework of the classical metaphysical and dialectical-idealistic tradition. It is believed that the modern form and meaning of the term civilization in some cases preserves what was put into circulation, first of all, by the French and English enlighteners in the second half of the 18th century, who proposed a universal scheme for the development of mankind - savagery, barbarism, civilization and focused on the search for the source of the development of society, the driving forces of history: V. Voltaire, J. Kondors, A. Turgot. This problem has not been developed in Uzbekistan from a philosophical point of view, in particular, that part of it that deals with ontological and epistemological aspects. In connection with the study, dialectical, formal-logical, historical-morphological methods were used, as well as comparative analysis methods.

1. Introduction

The changes taking place in modern Uzbekistan, the search for a new socio-economic orientation, have highlighted the question of its civilization and ways of its further development. Speaking without exaggeration, "civilization" is one of the main categories of philosophy, and in particular the part that deals with its ontological and epistemological aspects of its development, in addition, it is obvious that this category is one of the significant phenomena of social reality. Each civilization has a number of distinctive properties that have a huge impact on its history, political and socio-economic structure. The disclosure of their own characteristics is extremely relevant for modern Uzbekistan. It is important to analyze the prospects for the development of Uzbekistan in the context of inter-civilization relations against the backdrop of global changes in the world.

2. Explanation of thematic literature

It is important to note that the analysis of the formation of the concept of "civilization" appears on the pages of the works of the Marquis de Mirabeau, "Friend of men or a treatise on population" and "Friend of women, or a treatise on civilization"; in Ferguson's work, "An Essay on the History of Civil Society," where civilization is interpreted as the highest stage in the development of society compared to barbarism[1]. S. Montesquieu in his work "On the Spirit of Laws" identifies two stages in the life of mankind and explains the need for a transition to a civilized existence. For S. Montesquieu, civilization is a society where they are engaged in agriculture, live in cities, have money and have a civil code of laws [2].

The contrast of civilization and culture marked the thoughts of N. A. Berdyaev, A. Toynbee, O.Spengler [3]-[5].

* Djurayeva Nigora Avazovna, nigoradjurayeva@mail.ru

In the works of M.A. Barg, N.Ya. Bromley, K. Marx, L. Morgan, F. Engels, civilization is considered as the level of development of society "based on the upward movement" arising in the course of the dialectical interaction of material productive forces and production relations [6]-[8].

Coverage of the cultural and historical foundations of civilization is given in the works of M. Weber, N. Ya. Danilevsky, P. Sorokin, A. Toynbee.

On the basis of this material, a number of modern Russian scientists: B.Erasov, L.Novikova, L.Reisner - reveal the significance of the civilizational structure of society, based on the general principles of organizing people's lives.

3. Materials and methods

In connection with the study, dialectical, formal-logical, historical-morphological methods were used, as well as comparative analysis methods.

4. Analysis and results

The paradigm of civilization is a manifestation of the aggregate set of knowledge that affirms the specific system of its construction and shows the optimal mechanism of its existence, but it is relative.

When analyzing the ontological-epistemological problems of studying the development of Uzbekistan in the context of globalization from the point of view of the paradigm of civilizations, it is important to pay attention to the fact that the main content of the last decade of the twentieth century was the search for a new paradigm of world structure, which also affected Uzbekistan. In the inter-paradigmatic decade, the theory of modernization was again in demand. Nevertheless, it should be noted that there could be no universal scheme or model of modernization. There are no ready-made, the same for all cultures and civilizations programs for the reconstruction of society. Modernization is a modern stage of the socio-historical, ontological-epistemological processes that exclude a single and linear development path.

It is the presence of political independence as a prerequisite for the development of an independent civilization that can be considered another principle for the formation of a paradigm of civilizations. There is not a single civilization that would achieve a high degree of maturity without political independence. Without it, the emerging civilization turns into a tool that serves the interests and goals of others.

Modernization, by contrast, strengthens national cultures that are becoming more modern and less western. Peoples and generations are changing, scenes of historical life are moving, social orders are changing, civilizations are alternating, but the thread of historical development is not interrupted. And generation after generation preserves the cultural stock that managed to survive in the stormy historical ocean. Modernization is turning into a panacea, miracles are expected from it, all hopes for a bright future for the Uzbek sociocultural system are associated with it. On the one hand, in the conditions of an unprecedented technological breakthrough made recently by leading Western countries, as well as Asian states that are gaining strength, a bet on modernization is not only desirable, but also necessary. However, it is important to emphasize that modernization, like any other socio-cultural process, does not take place on its own, its effective implementation is preceded by a number of preparatory measures, the main of which is the determination of the goals and objectives of modernization, the ways and means of its implementation, sociocultural and civilizational consequences of modernization transformations.

Globalization, the new openness of the world is a challenge, but a positive challenge that we must be able to use in order not to stay on the sidelines among the laggards. Our task is not to fence ourselves off from globalization processes, but to develop a comprehensive strategy for adapting to new processes and phenomena that is beneficial for society and the country [9].

Today in Uzbekistan, the problems of stabilization and reform of various spheres of human activity are being solved based on our own development model, and our own programmatic approaches have been developed. The sociocultural features of Uzbek modernization are it catching up, fragmented, incomplete nature. In Uzbekistan, the Head of State Sh.Mirziyoyev developed an Action Strategy for 2017-2021, the purpose of which is to radically increase the effectiveness of the reforms, create conditions to ensure the comprehensive and accelerated development of the state and society, modernize the country and liberalize all spheres of life. In particular, five priority areas of the country's development have been identified:

- Improving state and social construction;
- Ensuring the rule of law and further reform of the judicial system;
- Development and liberalization of the economy;
- Development of the social sphere;
- Ensuring security, interethnic harmony and religious tolerance, implementing a balanced, mutually beneficial and constructive foreign policy.

The own path of our republic is based on the centuries-old experience of developed countries, as well as on the comprehensive consideration of the national historical heritage, lifestyle, traditions and mentality of the people of Uzbekistan.

It is important to note that the implementation of the Action Strategy will be a powerful impetus for the progressive movement of the Republic of Uzbekistan along the path of reforming and modernizing the country, building a legal and democratic state with a developed market economy, a strong civil society, ensuring the rule of law, security and the rule of law, inviolability of state borders, and interethnic consent and religious tolerance in society.

In the ontological-epistemological study of the paradigms of civilizations in general, as well as the development of Uzbekistan in the context of globalization, the role of the synergetic approach in philosophy is important. The specificity of the synergetic approach to the study of the development of the paradigm of civilization lies in such a modeling of the dynamics of civilization processes that takes into account the interconnectedness of ordered and chaotic states of the system, the variability of the ways of its development and the fundamental inevitability of the prognostic horizon, which is usually very close to the current state of the system. The value-semantic content penetrates into the synergetic approach mainly under the guise of improving the adaptive abilities of society to its increasing capabilities. The synergetic approach allows you to integrate various theories of civilization, in conceptual terms it can be used as one of the hermeneutics of the process of world-historical development of

civilization, and in terms of modeling civilization processes lasting several decades, it significantly complements already known concepts, changing the very structure of the study.

It is necessary to pay attention to the fact that from the point of view of development, every civilization or culture has its own hidden code, which includes a set of rules or principles that permeate all its inherent types of activity[6]. For example, the basic principles of industrial civilization have become standardization, specialization, synchronization, concentration, maximization and centralization. Moreover, at the heart of civilization are three types of backbone ties:

- spatial, extending to the features of the natural environment, landscape, which can also be called a geographical factor in the development of the paradigm of civilizations;
- temporary, forming a community of historical being, traditions, widely understood culture, language.
- social, cementing super ethnic community through the formation of comparable value-normative mechanisms, subjective awareness of community, self-identification of people.

Sociocultural codes, a combination of which identifies the essence of Uzbek civilization, its difference from others, can be attributed to:

1. Historical heritage, which is a temporary backbone link. Due to its unique geographical location, Uzbekistan has rich historical roots. The territory of the country became an eyewitness and experienced the change of many cultures and religions during various periods of the region's development, from the conquests of Alexander the Great to the campaigns of Genghis Khan. The formation of the Uzbek nationality is closely connected with the ancient peoples of Central Asia - Sogdians, Bactrians, Sako-Massageti and other tribes that for centuries inhabited the Central Asian rivers Amu Darya and Syr Darya and neighboring areas. Here were the centers of many powerful empires, such as the empire of Amir Temur. However, the name itself - the Uzbeks - was established only in the late 15th - early 16th centuries with the advent of the Sheibanid dynasty in Maverannahr, but the very concept of "Uzbek" was not used as a separate historically cultural type. Only in the Soviet period did this concept begin to be used in the context of statehood. Such a bright period of historical development largely explains the accumulated great spiritual and cultural potential.
2. National tolerance based on the Asian nature of peoples is a feature inextricably linked with the previous one. Uzbekistan, which has absorbed tribes and nationalities, has retained them as much as it received. Dozens of nationalities and nationalities live here, including residents of the Central Asian region: Uzbeks, Karakalpaks, Tajiks, Turkmens, Kazakhs, Kyrgyz, Uighurs, Dungans; Western and Eastern Slavs: Russians, Ukrainians, Belarusians, Poles; the numerous diasporas in Uzbekistan are Koreans, Iranians, Armenians, Georgians, Azerbaijanis, Tatars, Bashkirs, Germans, Jews, Lithuanians, Greeks, Turks and many other nationalities. Such ethnic diversity is due to various historical events. Many representatives of the indigenous peoples of the Union republics of the USSR were evacuated to Uzbekistan during the Second World War (Russians, Tatars, Armenians, Belarusians, Ukrainians, Germans, Jews, etc.). Representatives of individual nations were deported from places of permanent residence during the years of Stalinist repression (Koreans, Crimean Tatars, Chechens and others).
3. Tolerance. Scientific sources indicate that Uzbekistan is one of the most ancient places in the world where the first religious views and ideas arose. This is eloquently evidenced by the ancient burial places in the Teshik-Tash caves, the bas-relief finds of Faez and Kara-Tepa, the finds in Dalverzintepa and Sogdiana, the records of Darius the first on the Behistun rock, as well as the results of studies conducted jointly with French and Japanese archaeologists. Since ancient times, Zoroastrianism, Buddhism, Judaism existed in our region, which then turned into complex ideological systems. On the eve of the Arab offensive and the emergence of the religion of Islam in the region, Nestorian Christianity penetrated here from Syria through Iran and took an important position. At the beginning of the VI century, a Nestorian bishop served in Samarkand, and in the VIII century - Nestorian Metropolitan. In addition, in ancient times, the population professing Christianity lived on the territory of the present Tashkent and Khorezm regions. In the X century A.Hanafi's school emerged in the territory of the modern Uzbekistan region, which is distinguished by its tolerance that has formed at the moment, this was not an accidental phenomenon. The culture of Uzbekistan, which absorbed the best achievements of the past, had, in turn, a tangible impact not only on the subsequent development of Islamic culture, but also on the Western Renaissance, on the development of principles of tolerance on a global scale. With the spread of Islam in Uzbekistan, a meaningful dominant of the spiritual space of the state is being established. Apparently, it should be agreed that it was Islam as a cultural dominant that formed the basis of the civilizational development of the Uzbek people and largely extent formed the national spirit of society, the metaphysical reality existing in Uzbekistan. Over the years of independence of Uzbekistan, religious education has risen to a qualitatively new level. Today, we can confidently say that we have formed a complete system of religious education, which includes the Tashkent Islamic Institute, ten specialized secondary Islamic educational institutions (madrassas), Orthodox and Protestant seminaries. The cooperation, commonwealth, interethnic harmony and religious tolerance that have existed for centuries between the peoples of our region acquired a peculiar content during the years of independence. This also confirms the same attitude on the territory of our multinational and multiconfessional country to every citizen of our state, regardless of race, nation, religion and mother tongue.
4. Spirituality. Uzbekistan is characterized by a special spiritual disposition, characterized by the predominance of the heart over will, contemplation over analysis, conscience over reason, prudence and practical sobriety, freedom and will overpower and coercion, compassion over personal gain and self-interest. These features, in their totality incomprehensible to representatives of other civilizations, have led to a special spiritual depth, which is

reflected in our music, painting, literature, culture as a whole.

5. Collectivism. The dominance of collectivism as a value is confirmed by all cross-country comparative studies. The whole history of the development of our civilization - this is the history of the survival of peoples in natural and climatic conditions, the development of uninhabited spaces, the protection of borders, the vital need for mutual assistance, in collective work - have shaped community, collectivism as a hallmark of mentality.
6. The Uzbek language. For the development and functioning of civilization, the functions of language as a means of consolidating the community, a means of ensuring social interaction, and a means of accumulating and transmitting social experience are especially significant.

In historical terms, the civilizational activity of society contributes to the formation of national languages in the course of a series of convergent-divergent processes: the growth of social integration leads to the strengthening of the speech unity of society, and the increasing differentiation of social life determines the internal differentiation and enrichment of the language system, and the expansion of its communicative functions. Historically important is the emergence of nation-states and urban settlements that support centripetal forces in society and ensure social stability - one of the main prerequisites for creating a national language.

The ability to communicate is one of the factors satisfying human needs and is the paradigm foundation of civilizational existence. Language is not just a means for exchanging thoughts, but also a kind of spiritual authority, which reflects the nature and history of the people, their spiritual traits, mentality, and temperament. It is with the help of language that cultural values are exchanged, the spiritual integration of civilization takes place. Thanks to the emergence of alphabetical writing and a single language, knowledge takes shape and becomes more accessible, accumulates and takes shape in the form of scientific theories of knowledge, develops common moral standards that contribute to the process of self-identification and the growth of national self-awareness. We can agree that "civilization is the largest" "we" "... that breaks up into a huge number of smaller ones: identification with a country, region, with a local community" [10].

M.Barg believes that "the emergence of civilization became possible after the advent of writing" [6]. The development of writing and number systems is becoming necessary in administrative affairs of complex political systems, in the management of large lands and in commercial affairs. The introduction of writing to replace mnemonic methods had revolutionary consequences, since it was used to satisfy a wide range of new society needs. It was used to write laws, record cosmological knowledge, dynastic stories, to register transactions, as well as to write magic formulas, which was of no small importance in early civilizations. Writing refers to the functionally significant criteria of civilization. Subsequently, writing acts as an objective form of self-awareness.

In particular, with the help of writing, a person got the opportunity to communicate with people who are far beyond the reach of his voice, to control vast populated territories. In addition, writing allowed a person to go beyond the present. Thanks to it, the opportunity was opened not only to know the past, but also to turn to the future. As Giddens notes, writing, "expands the boundaries of spatio-temporal distance and creates a perspective of the past, in which the reflective comprehension of newfound knowledge can be separated from the established tradition" [11].

It should be noted that the concept of "civilization" in the singular and today retains its significance when we discuss the problems of global civilization that affect all of humanity. These are, first, the problems of environmental and nuclear safety, the preservation of humanity, the fight against epidemics and others. Turning to the problems of civilization gradually leads to the formation in science of various directions or approaches that are important for a thorough understanding of the studied sociocultural phenomenon of civilization.

As mentioned above, from the point of view of synergetics, the development of the paradigm of civilizations is based on system-forming connections, some of which are called spatial, they form the geographical factor of the paradigm. Social laws, the way of life of peoples, their social system, mores correspond to physical conditions, climate, soil properties, geographical location of the country. The political system of the state, the form of government also depends on the territorial size of the state. L.Mechnikov in the work "Civilization and the Great Historic Rivers" carefully traces the connection between the distribution of mountains and rivers, and so on with the development of countries and nationalities. He writes: "Is it possible to create any generalization of the cultural and historical meanings of the geographical environment in all their diversity ... work on the discovery of a formula that encompasses in general terms those hidden relationships that bring together and link each phase of social evolution, each period of history humanity with a certain state of the geographical environment" [12]. Representatives of the geographical approach note that primary civilizations arose in an ecologically favorable subtropical climate. Consequently, favorable natural and climatic conditions, the availability of water resources allowed mankind to move to a productive economy. Moving to the subtropics, as well as territorial expansion, has a direct impact on the occupation of the population, and determines its economic and economic base. The uniqueness of natural conditions affects the psychology of the population, the nature of culture, contributing to the formation of its specificity and identity.

Bearing in mind the arguments of the proponents of a geographical approach, we agree that for a comprehensive explanation of the concept of "civilization", the geographical factor is important. In our opinion, the geographical factor has a huge impact on the development of the paradigm of civilizations, but it cannot be considered absolute, since civilization is not only a territorial space formed as a result of the influence of the climatic conditions of a particular region. The natural environment for civilization is that element that must first be understood and then overcome. Accordingly, the development of civilization will depend on the innovative activity of society as a response to the challenge of nature. Uzbekistan is the largest state in Central Asia. In the past, the country had a very favorable location for trade, being on the ancient Silk Road from Europe to the Far East. If we consider the geography of Uzbekistan in general, it can be noted that the vast majority of the territory are deserts and mountains, valleys and lowlands occupy only a small part. At the same time, almost 80 percent of Uzbekistan is the Kyzylkum desert. The southeastern and northeastern parts are represented by mountainous terrain, the elements of which are the foothills and lower mountains of the Tien Shan system. Due to the hot and dry climate, the predominance of mountain and desert

landscapes in Uzbekistan, residents settled unevenly across the territory. The population is concentrated mainly in oases. In the desert regions of the republic, the population density is very low. For example, in Karakalpakstan and the Navoi region, there are only 7-9 people per square kilometer of the territory, while in the most densely populated region of Uzbekistan - the Ferghana Valley - about 500 people per square kilometer. This is the highest indicator of population density not only among the CIS countries, but also one of the highest in the world.

There is one more systemically important factor in the framework of the synergistic approach to the development of the paradigm of civilizations - axiological, whose supporters in the study of the paradigm of civilizations consider values to be the basis for dividing into different civilizations, which are defined as the expression of the general attitudes of their time. Values determine not only spiritual culture, but also material, therefore, cover all aspects of society. Values pervade the whole life of society, form people's attitude to the world, faith and turn into conscious or unconscious psychological attitudes. We can say that values to a greater extent influence the creation of a national-ethnic community, where value guidelines are formed in the process of cultural development. Values provide an opportunity for a qualitative change in society, both in the spiritual and social spheres. Here, it should be concluded that in the value sphere there is a curious phenomenon: with the advent of new urgent problems and new values corresponding to these problems, the old attitudes are partially revised and are gradually being rebuilt in the spirit of the times. For example, quite radical changes in the value sphere were accompanied by industrial revolutions in the West and in the East. This suggests that the value systems of various civilizations in the future will come closer to certain limits as the field of general problems and common moral norms expand. Already today, the main principle of the new positioning of values in various civilizations is the desire to avoid the risk of environmental catastrophe and exacerbation of global problems. However, fundamental civilizational values persist.

In Uzbekistan, these values include the followings:

- Vital values: the right to life and procreation, the preservation of health, the integrity of the person, as indicated in the Basic Law of the state;
- Environmental values: clean soil, water, air, the adequacy of basic resources.
- Primary civil rights: protection of unlawful violence and coercion, freedom of movement, inviolability of the home, freedom of conscience, speech, assembly, association.
- Universal political and legal values: independence of the court, freedom and independence of the press, various forms of citizen participation in political life
- Universal socio-economic values: the right of everyone to self-sufficiency, the right to work and others.

All of the listed values serve as the basis for the implementation, firstly, of civilizational values proper, related to the peculiarities of the lifestyle and thinking of each civilization, and secondly, of national values associated with the characteristic features of each of the peoples included in this civilization. The recognition of universal civilizational values, aimed at protecting civilizational and national values, as leading and defining, today constitutes the minimum platform that is necessary for consensus in the dialogue of cultures.

When considering the most significant approaches to understanding the essence of civilization, it should be noted that the above descriptions and definitions are characteristic of a comprehensive understanding of civilization as a complex, developed society in which both economic factors and the social system, as the moral principles of the regulation of relations, are important, and the political structure, both practical knowledge and aesthetic ideals. We define civilization as a societal cultural-historical system, uniting at the superethnic level a population with a common self-identification, organically connected with a specific natural and spatial environment.

It is also important to note that the process of development of the paradigm of civilization on a macro-chronological scale is represented by the concepts of “moral synergetics”. The study of the paradigm of civilization from the standpoint of self-development of the method of material production focuses on the alternation of the regular stages of its evolution and transition periods containing alternative paths of civilizational development. “Moral” synergetics, however, emphasizes that the transition of mankind from a natural state to a civilized one and the subsequent development of civilization were mainly associated with the creation, selection and subsequent consolidation of the most effective mechanisms for maintaining the sociocultural integrity of society[6].

Some researchers of the development of civilizations have identified attractors that draw it into the lumen of cultural and historical life. So, for example, N.Danilevsky acted as an attractor with an original national idea, Spengler - “prasymbol”, for Toynbee - God, for Kreber - style, for McNeill - text, for S.Huntington - culture. It can be said that in the globalizing conditions of development for Uzbekistan, such attractors of development as the national idea, symbolism, religion, language and writing, original culture and millennia-old traditions are important and there are.

However, the trajectory of civilization is not limited to the upward movement, it also includes a number of other phases. In the initial phase of the formation of civilization, its slow growth occurs. It is being replaced by a phase of accelerated, exponential growth. In this phase, the trajectory of civilization rises most rapidly. However, accelerated growth cannot continue continuously. Encountering internal and external restrictions, it is inevitably accompanied by an increase in contradictions. The growth of such contradictions indicates that civilization has entered the critical zone of its development. It is during this period that she faces a challenge from the environment. The challenge requires the resolution of urgent contradictions, and depending on how successful the answer is, various paths for the further development of civilization are possible.

The critical zone is, in fact, a region of bifurcations where the branching of the trajectory of civilization takes place. Moreover, in our opinion, three main ways of its further changes are possible. Firstly, civilization may fail, which, as a rule, is accompanied by the collapse of the previous structure, a decrease in the level of complexity and a transition to another, much lower trajectory of movement. Secondly, civilization can go through its entire life cycle, including a period of upward changes, a “plateau” and a downward phase of movement. Finally, a third option is possible, in which, having found an adequate response to the environmental challenge, having overcome the ascending section in its development, having passed the “plateau”, civilization can again

continue the upward movement. In the development of civilization, not only its breakdown may arise, but also its disintegration, or transformation. The era of the breakdown should be considered complete when, as a result of the promotion of this mechanism, a return to the modification of the nucleus or at least the forcible introduction of new stagnation is no longer possible. This border is conditional, and then the same processes of breakdown unfold in the collapse of civilization or transformation into other civilizations.

5. Result and Discussions

In this regard, it is also important to emphasize the threats that threaten the existence of civilizations and cultures:

- the "tension" generated by the conflict of the "organic" social needs of man and the abstract social requirements of civilization;
- the inability of society to find ways of sociocultural adaptation to its own growing opportunities;
- inconsistencies in the development of the economic, social and informational subsystems of civilization.
- from the excess pressure exerted by it on human nature and the environment, the mismatch between the demographic and sociocultural spheres of the development of civilization.

As modern mankind has approached the threshold of permissible anthropogenic pressure on the biosphere, relations between society and the environment have become sources of civilizational conflicts. New natural restrictions require a change in the development strategy of mankind in general and individual civilizations in particular. The Moiseev are offered two key directions for changing this strategy, which do not go beyond the circle of recipes outlined earlier by the Club of Rome: technical and technological re-equipment of mankind and the adoption of a new morality, both of which are dependent on the "environmental imperative" [13]. Synergetics has established that stabilization of a nonequilibrium state is possible only due to an increase in entropy in other systems, an increase in chaos at lower hierarchical levels of being. Consequently, environmental and man-made crises in relation to the development of the paradigm of civilization do not have an eventual but an attributive status. And if the technological and economic progress of civilization is not accompanied by a corresponding improvement in the institutions of sociocultural integration, then no technology for adapting society to the environment will help him overcome the barrier of another ecological crisis.

In the future, the development of civilization:

- will lead to the creation of an ecumenical system;
- needs to expand its own technical basis;
- will be determined mainly by innovations in the information subsystem.
- can lead to the creation of the noosphere, when the causal relationship between nature and civilization will reverse and civilization will become a factor in the development of nature.

6. Conclusion

Making a conclusion, it is important emphasize that in the present work, the idea of the uniqueness of each civilization and culture is developed, and this does not allow building their hierarchy, gradation "above" and "below". Social norms, values, factors, principles of organization, social institutions of one civilization are not standard for others. The features and results of modernization process are always determined by the previous development and predetermined by the cultural, historical, social features of the country. The modernization process is determined by history, which means a conceptual limitation of the mass of possible solutions at each stage of development. The unlimited borrowing of foreign cultural values, institutions leads to the loss of certainty, the loss of one's own path of development, degradation and destruction of the system, and the violation of the intrasystem determinism of values.

Uzbek society is a special world, civilization, therefore, it is necessary to analyze it, exploring the corresponding phenomena and categories inherent in it. It is a special social organism, distinctively isolated from other societies and having common spiritual sources and foundations, common territory, common economic and geopolitical space.

References

[1] Benvenist E. General Linguistics. Civilization. Word history. –M., 1974. https://platona.net/load/knigi_po_filosofii/filosofija_jazyka/benvenist_ehmil_obshhaja_lingvistika_pod_red_ju_s_stepanova_izd_1974_g/32-1-0-708. -P.448.

[2] Montesquieu S. On the spirit of laws // Elect. production - M., 1955. –P.357. URL: http://www.civisbook.ru/files/File/Monteskye.O dukhe.pdf

[3] Berdyaev N.A. The philosophy of free spirit. –M., 1994; ISBN 5-250-02453-X. DOI: 10.24045/conf.2017.1.16

[4] Toynbee A. Comprehension of history. –M., 1991; ISBN: 978-5-8112-3798-2.DOI: https://doi.org/10.14258/izvasu(2015)3.2-39

[5] Spengler O. Sunset of Europe. –M., 1993. DOI: 10.17805/zpu.2016.1.2. –P.663.

[6] Barg M.A. A civilizational approach to history: a tribute to conjuncture or a requirement of science? // Civilizations. –M., 1993. –Vyp.2 .; P.46. https://cyberleninka.ru/article/n/ponyatie-tsivilizatsionnyy-podhod-v-otechestvennoy-istoriografii-na-rubezhe-xx-xxi-vekov-1

[7] Barg M.A. Eras and ideas. –M., 1987 .; https://imwerden.de/pdf/barg_epokhi_i_idei_1987.pdf -P.348.

[8] Bromley N.Ya. To the question of the relationship between the concepts of "civilization" and "formation" // Civilization. –M., 1992. –Vyp. one.;

[9] Karaganov S.A. Chances and threats of the new world // Russia in global politics. 2003. No2. –P.13-14. https://globalaffairs.ru/number/n_771

[10] Orlova I.B. Civilization paradigm in the study of the socio-historical process. Dis. On the co. Uch. Step. Doc. –M., 1999. –P.167.

[11] Morgan L. Ancient society. –L., 1934; -P.104. http://www.pseudology.org/Sex/MorganLG_DrevneeObshestvo2.pdf

[12] Mechnikov L.I. Civilization and the great historical rivers. –M., 1995. –P.81. ISBN: 978-5-8112-5112-4 https://cyberleninka.ru/article/n/tsivilizatsiya-i-velikie-istoricheskie-reki-geograficheskaya-teoriya-progressa-i-sotsialnogo-razvitiya

[13] Moiseev N.N. Modern anthropogenesis and civilizational faults // Problems of Philosophy. -1995. -№1. P.6, 24 .;

[14] Weber M. Selected Works. –M., 1990 .; ISBN 5-01-001584-6 –P.808. https://www.gumer.info/bibliotek_Buks/Sociolog/vebizbr/index.php

[15] Condorcet J. Sketch of the historical picture of the progress of the human mind. –M., 1936 .;

[16] Erasov B.S. Culture, Religion, and Civilization in the East: Essays on the General Theory. –M., 1990; ISBN 5-02-017165-4. –P.205. https://search.rsl.ru/ru/record/01001569098

[17] A new post-industrial wave in the West. Anthology / ed. V.L. Inozemtsev. –M .: Academia, 1999. –P.104. ISBN 5-87444-067-4. http://iir-mp.narod.ru/books/inozemcev/page_1003.html -P.640.

[18] Novikova L.I. Civilization and culture in the historical process // Questions of philosophy. -1982. -№10 .; Novikova L.I. Civilization as an idea and as an explanatory principle of the historical process // Civilizations.-M., 1992. –Vyp.1 .;

[19] Peccechi A. Human qualities. –M .: Progress, 1985. –P. 312. https://www.twirpx.com/file/359138/
[20] Reisner L.I. Civilization and the way of communication. –M., 1993. ISBN 5-02-017681-8 . https://search.rsl.ru/ru/record/01001676355 –P.304.
[21] Ruzavin G.I. Self-organization and organization in the development of society// Questions of philosophy. -1995.-№8 10.7256/2306-0158.2013.4.581 URL: https://nbpublish.com/library_read_article.php?id=581. -P.71.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 34-40 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Dynamic Objects Parameter Estimation Program for ARM Processors Based Adaptive Controllers

Vasiliy Olonichev*, Boris Staroverov, Maxim Smirnov

Department of Automation and Microprocessor Technology, Federal State budgetary Educational Institution of Higher Education Kostroma State University, Kostroma, Russia

ARTICLE INFO

Article history:
Received: 08 January, 2020
Accepted: 22 April, 2020
Online: 03 May, 2020

Keywords
Adaptive control
LSQ parameter estimation
ARM_CORTEX processors
Program optimization

ABSTRACT

Modern microcontrollers are capable to realize not only traditional PID-regulators but also adaptive ones. Object of control parameters estimation is the biggest part of adaptive control from the point of view of time consumption. The ways to reduce this time for digital control systems based on ARM-CORTEX 32-bit and 64-bit processors are shown in the article. These ways include source code refactoring, using vector registers and parallelism of code. As result of program improvement, a new algorithm for least squares method was suggested. Intrinsics for vector operations and OMP directives were added to the program to realize data and code parallelism. All options were tested for time consumption in order to find out the best decision. The program suggested may be useful while realizing adaptive controller based on single-board mini-computers and microcontrollers

1. Introduction

This work is an extension of conference paper “Optimization of the Program for Run Time Parametrical Identification for ARM Cortex Processors” originally presented in "2018 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)" [1]. Conference paper has the results obtained only with the 32-bit armv7 processor. The results obtained with the 64-bit armv8 processor are added in the current paper. Also, the questions of alignment data in memory and leftovers processing during vectorization are considered.

Nowadays most of the micro-controllers are based on inexpensive but at the same time powerful ARM processors. High computing abilities of these processors allow to realize not only simple PID-controllers but more sophisticated adaptive controllers. Adaptive digital controllers are indispensable for technological processes which require high quality of control. In this case oscillation and overshooting are inadmissible and setting time must be minimal. And digital controllers are able to improve process control performance significantly. The theory of digital control systems was developed in 70-80 years of the last century. In particular, K. Astrom and B. Wittenmark [2] and R. Iserman [3] showed that digital controllers are the best when aperiodic transient processes are wanted and described how to make state variable modal digital controller capable to provide any in advance known characteristics of the transient process. When an object of control is timeinvariant it is possible to use experimental data and find out the coefficients of the digital transfer function of this object and the coefficients of digital modal controller and digital observer preliminarily. These calculations are carried out only once and their results may be used as the constants in the program for direct digital control. But when the object of control is unstable i.e. its characteristics drift with the time, or its characteristics are non-linear and its linear approximation depends on operating point, all above mentioned calculations must be made repeatedly at run time within controller itself with the pace of technological process. And in this case time consumption for such calculations may be crucial. In other case quality of control may decrease drastically.

2. Using least squares method for parameter evaluation

The most substantial part of calculations in discussion is the object's parameter estimation, i.e. the process of finding out its digital transfer function coefficients. And the task of this article is to show how time consumption of the program for parameter estimation may be reduced. The novelty of this work is in making new program realization of the least-squares method that includes refusing of function decomposition of the code. As a result, some intermediate matrices may be dropped and outer loops of sequential stages of calculations may be linked. Such a decision allows reducing the program's time and memory consumption that is very significant in the case of real-time applications based on the microcontrollers. Also, for further optimization, the code

*Corresponding Author: Vasiliy Olonichev, v_olonichev@ksu.edu.ru

parallelism and different ways of data parallelism are added to the program and a comparison of the obtained results is made.

The general method of the object of control parameter estimation is the least-squares method or LSQ [4]. Using LSQ, one can get a transient function of the object in a discrete form. The coefficients of this transfer function are used to find out parameters for tuning adaptive dynamic regulators [3].

Despite the fact that LSQ is a rather old method, there are still a lot of publications connected with it. Each of them is devoted to special ussue. For example, identification of nonlinear systems [5], time-varing magnetic field analysis [6], signal processing [7], lines approximation in multidimentional space [8]. But among them, researches connected with LSQ utilization at microcontrollers when amount of RAM is limited and time constraint are strict are not found.

LSQ is widely used in many fields of applied calculations and it's realization is available in many program libraries. First of all must be mentioned lapack library (https://www.netlib.org/lapack/) - standard library that may be found in many Unix-like operating systems. Another widely used library is gnu scientific library or gsl (https://www.gnu.org/software/gsl/). These libraries use QR-decomposition [9] for solving LSQ problem. This method is considered to be one of the fastest for the big dimension tasks.

But the objects of control usually are described with the models that have order between 2 and 6. So it is reasonable to suggest that program based on QR-decomposition will spend more time making preliminary computing than useful one when applied to such small systems. And with the micro-controllers situation when it will be not enough memory for such libraries is also possible.

To check this hypothesis the program was written that applied matrix operations as LSQ suggests [9] and as is shown in (1):

$$\begin{bmatrix} A \\ B \end{bmatrix} = \left[X^T \cdot X\right]^{-1} \cdot X^T \cdot Y \tag{1}$$

where X – matrix with the dimension of Nx2M filled with values taken from input and output of the object with the regular intervals of time; Y – vector with the dimension of N filled with the values taken from the output of the object; A – output vector with the dimension of M having coefficients of transfer function denominator; B – output vector with the dimension of M having coefficients of transfer function nominator; L – number of experimental points; M – order of the object of control; N = L – M – number of equations.

Covariance matrix is represented in (2):

$$C = \left[X^T \cdot X\right]^{-1} \tag{2}$$

Equation (2) that is the part of (1) plays special role and is used for evaluation of parameters dispersion [10].

How to fill matrix X and vector Y with the experimental values is shown in [3].

Other reasons to write such program are the following:

At the first, for correct parameter estimation of the object of control the trace of covariance matrix must be minimal and the value of this trace depends mainly on sampling period [11]. This means that at the first stage of parameter estimation time step may vary and alongside with the model's parameters the trace of covariance matrix must be computed. Getting the covariance matrix is the part of LSQ method, so it's trace may be computed when direct matrix operations are involved. While using library functions this data are hidden inside them. And to get covariance matrix one needs to repeat a bigger half of computations already made.

At the second, during LSQ parameter evaluation some dynamically allocated matrices are used for storing intermediate data. Each time when library function is called the memory is allocated for them and then released. While using our own code we can allocate memory for intermediate matrices only once and then use them with every next time step.

As equation (1) shows LSQ consists of matrices multiplication and matrix inversion. Matrices multiplication may be written from scratch. An example of matrix inversion can be found in the Internet (http://www.programming-techniques.com/2011/09/numerical-methods-inverse-of-nxn-matrix.html). This program realizes the Gauss-Jordan method [12].

3. Program Testing and Optimization

3.1. General Information about Testing

For realization and testing of the programs singleboard mini-computers CubieBoard-3 and Odroid-C2 were used. CubieBoard-3 has two core CPU CORTEX-A7 (ARMv7) with the frequency 1GHz and 2G of RAM. Operating system Linux Ubuntu 18.04.1 with the kernel 4.19.57 is installed on this computer alongside with gcc compiler v 7.4.0. Odroid-C2 has quad core CPU CORTEX-A53(ARMv8) with the frequency 1.5GHz and 2G of RAM, operating system Linux Ubuntu 16.04.09 with the kernel 3.14.79 and gcc compiler 5.4.0. For the conference paper [1] programs were made with the gcc v 4.6.3, so results presented in this article may slightly differ, first of all due to the fact that the realization of optimization in these compilers is not identical.

For program realization, the C++ language was chosen. It's a common practice nowadays even for embedded systems. If you do not use classes with the virtual functions, the productivity of the result code is almost the same, and at the same time, the full power of C++ as a language of generic programming is available.

All programs were compiled with the -O2 level of optimization.

Working with the matrices, one must decide how to store them in memory. First way suggests using dynamic one dimensional flat vectors. Matrix elements in this case are accessible with the function or overloaded operator () taking as arguments row and column number. E.g. getElem(X,i,j) or X(i,j). Second way suggests using dynamic two dimensional arrays. In this case X[i] is a vector of pointers containing the addresses of matrix's rows and X[i][j] is an element of matrix. First way is considered to be slightly faster as the data occupy continues space in memory. But for an adaptive controller second way is more suitable, because in this case matrix X must be renewed with every next step in time. Old data must be removed and new once added. With the flat vector all elements counted with thousands will be moved within

this vector from the end to the beginning. But with the two-dimensional array only pointers counted with tens will be moved.

3.2. Comparison of Code Using Library Functions and Code Using Direct Operstion with the Matrices

To compare time consumption for LSQ 3 programs were written. The first used function dgel from the lapack library, the second used function gsl_linalg_QR_lssolve from the gnu scientific library and the third realized equation (1) using self-written functions for matrix multiplication and inversion. Time consumption for the calculations was found out as the difference between the time measured before and after the calculations. To measure time the function clock_gettime was used.

Time consumption were determined with the matrices of the following dimensions MxN: 2x40, 3x60, 4x80, 5x100, 6x120, 7x140, 8x160 and 10x200. Objects of higher orders require more experimental points. But this does not mean that to evaluate parameters of the for example 4 order object, one must use exactly 80 experimental points. It is just an average value.

The matrix X and vector Y were filled with the random numbers because we need to get not the results of LSQ-evaluation but only time consumption for getting them.

Calculations for each program and for each dimension were repeated 5 times. Maximum and minimum were removed and out of the rest three measures, an average value was calculated. To run programs, calculate time consumption and to make plots the special script in Python language was written. The resulting plots with the time consumption against object order are shown in Figure 1 for CubieBoard3 and Figure 2 for Odroid-C2.

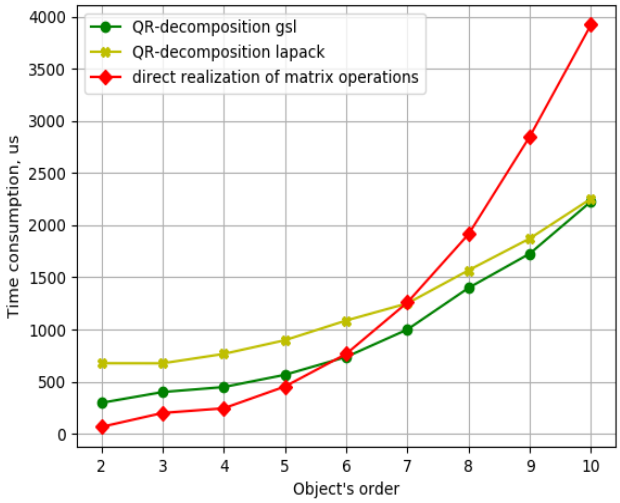


Figure 1: Time consumption for LSQ parameter estimation against object's order by the programs with the double precision numbers on CubieBoard-3.

The plots in Figures 1 and 2 show that the program that uses matrix operations takes less time then library functions for the tasks of small dimensions. But this program must be rearranged in order further to improve its efficiency.

First, decision must be made is there any sense to replace the double precision variables with the single precision variables or with the fixed-point ones.

3.3. Time consumption for Carring out Arithmetic Operations

For this purpose, it will be useful to find out how many time take arithmetic operations with the operands of the different types. And proper program with the three operands expression was written and run on CubieBoard-3 and Odroid-C2 computers:

a = b # c

where # in turn is + - * and /; b=2; c=3.

The results are shown in Table 1.

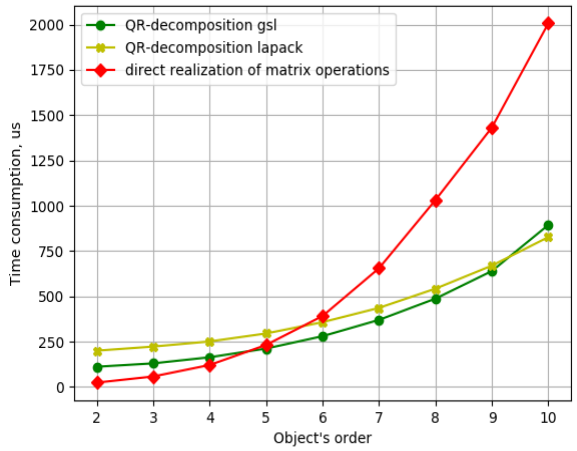


Figure: 2. Time consumption for LSQ parameter estimation against object's order by the programs with the double precision numbers on Odroid-C2.

Table 1: Time Consumption in μs for Carrying out 1000 Operations on Single-Board Computers CubieBoard 3 and Odroid-C2

Variable type	Operation			
	Addition	Subtraction	Multiplication	Division
CubieBoard-3				
int	13,5	13,5	14,6	30,2
long	13,5	13,5	14,6	30,3
float	14,6	14,6	14,6	29,2
double	14,6	14,6	17,7	43,8
Odroid-C2				
int	7,1	7,1	8,5	8,5
long	7,1	7,1	9,1	8,5
float	9,8	9,8	9,8	15,6
double	9,8	9,8	9,8	21,5

As Table 1 shows, the time consumption for processing integer variables is less than the time for processing floating point variables by 8% for armv7 and 28% for amrv8. But realization of LSQ in adaptive controller requires support of numbers in wide range of values. For example, parameter evaluation of the 3 order object using the experimental results where input varies between 0.0 and 1.0 and output between 0.0 and 200.0 will give values in the intermediate matrices varying from 1.0e-3 to 1.0e6. Obviously, integer values can't be used for such computing.

Time consumption for single precision and double precision floating point variables is identical for addition, subtraction and multiplication at both platforms. But division takes considerably more time especially for double precision variables.

3.4. Using Single Precision Variables and Parallelism of Code

Time consumption of the program depends not only on time required for the computing but also on cache misses [13]. Float variable takes 4 bytes and double takes 8 bytes. That means that the processor's cache can hold more data in case of float variables and so, cache misses will be met less often.

There is one more reason to use variables of single-precision:

- The vector unit of ARM-CORTEX-A (ARMv7) processes only single-precision floating-point numbers, and vectorization is a significant source of increasing program efficiency.
- Microcontrollers STM32 has hardware support only for single precision, and software emulation of double precision is rather slow.

Usually, it's recommended to avoid single precision variables in calculations [14]. But in our case, the dimension of the task is not big. And if to use in the experiments optimal sampling period, well-conditioned matrices will be obtained [11]. LSQ identification of the same object made with double precision and single precision numbers is almost identical. For these calculations, real experimental results obtained with the object with orders 3 and 4 were used.

There is also one more source of program efficiency improving. It is parallelism of code. And gcc compiler supports OpenMP specification, which provides parallelism or multi-threading. The results obtained are presented in Figures 3 and 4.

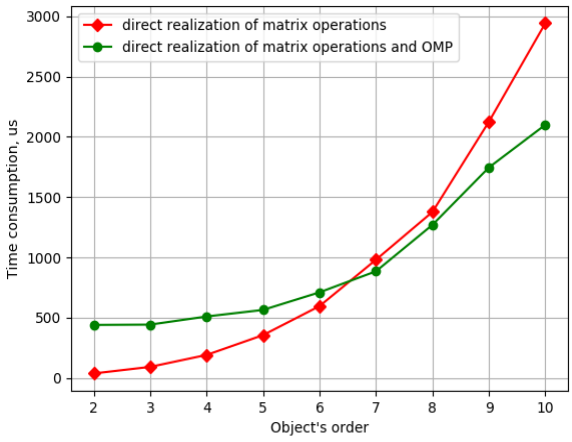


Figure 3. Time consumption for LSQ parameter estimation against object's order by the programs with the single precision numbers on CubieBoard-3.

The CubieBoard3 computer has 2 core processor and Odroid-C2 has 4 core processor.

Figures 3 and 4 show that using single precision variables instead of double has given about 25% increase in productivity. At the other hand on CubieBoard-3 parallelism has given expected results. I.e. while working with the small matrices synchronization between threads takes more time then parallelism saves it. And for objects with order 7 and higher parallel program becomes faster. But the results obtained on Odroid-C2 show that parallelism instead of increasing productivity reduces it in the whole range. The explanation of this fact may be following. The program calls several functions, and directives for parallel code are placed within them. So OMP preprocessor generates code that creates and cancels threads within each function. At CubieBoard-3 more modern gcc compiler was used with this problem fixed.



Figure 4. Time consumption for LSQ parameter estimation against object's order by the programs with the single precision numbers on Odroid-C2.

3.5. Refactoring of the Code

Next stage of optimization is concerned the code itself.

The first stage is also connected with the problem of cache-missing. In the programs written in C, matrices are allocated in memory row-wise. During multiplication one of the matrices is scanned row by row, and a big piece of data is loaded into cache. But another matrix is scanned column by column, and getting the next element may come to a cache-miss. If preliminary to transpose another matrix it will be also scanned row by row [13].

There is a division in the inner loops of the function that makes matrix inversion. If to calculate the inverse number in the outer loop and replace division with the multiplication we can get another source of the productivity raising.

From the point of view of the structural programming the code must be divided into functions each of them makes logically complete operation. In our case these are matrix multiplication and inversion. But its also known that such structural decomposition may reduce productivity of the program.

If to make one function that makes all calculations in one step it will be possible to take into account specific properties of computing. Informational matrix $X^T \cdot X$ is symmetric and it's possible to calculate only half of it. And also, it's possible to get away two intermediate matrices with the sizes [2*M][2*M]. And intermediate matrix [2*M][N] may be replaced with a vector with the size [N]. As a result, the number of cache-misses and total consumption of memory will be reduced. The last is especially significant for STM32 microcontrollers with limited RAM. In this one function, it's also possible to improve an algorithm. In this new program, the next stage of calculation will be started within an outer loop of the previous stage.

Alongside with the algorithm improvement, OMP directives were added to the code in order to apply parallelism. The resulting program was compiled twice. First time with an OMP flag to get parallel code. And second time without this flag to get code without multithreading.

Calculations were repeated with the same data. The results are presented in Figures 5 and 6.

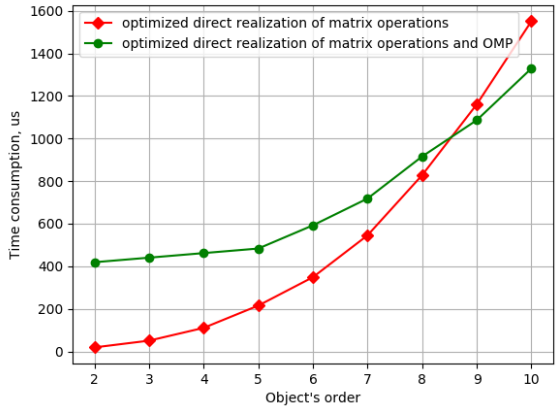


Figure 5: Time consumption for LSQ parameter estimation against object's order by the programs with the optimized function on CubieBoard-3.

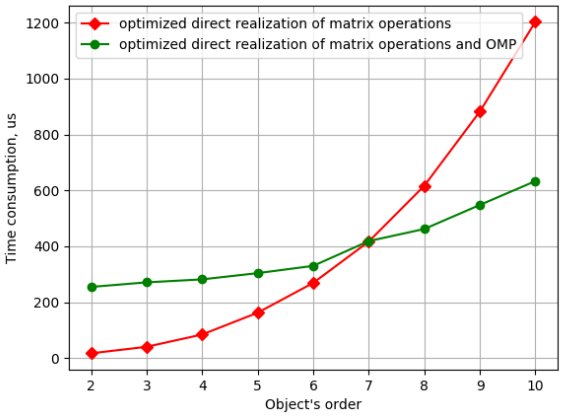


Figure 6: Time consumption for LSQ parameter estimation against object's order by the programs

Plots in Figures 5 and 6 show that optimization of code gave increasing in speed about 80% for CubieBoard-3 and 50% for Odroid-C2. And using parallelism in one function has given results on Odroid-C2. But also as one can see parallel code is faster only with 7 and higher-order object models. For the object's model with the order from 2 to 6 nonparallel code is faster and it may be recommended for practical utilization.

And more significant is the fact that the optimized program is faster than programs using lapack and gsl libraries in the whole range from 2 to 10 at both mini-computers.

3.6. Using Vector Operations

Vectorization is another way to increase program efficiency. ARM-CORTEX-A processors have the vector unit named NEON. This unit has 128-bit registers that enough to keep 4 floating-point numbers. The vector instructions process all numbers in the vectors at one time. Theoretically, it can increase efficiency by 4 times. But usually, this value is less because switching processor's pipeline between vector and regular registers takes a lot of time. (https://developer.arm.com/products/processors/cortex-m/).

There are many options to use vectorization: special libraries, auto-vectorization of compiler, OMP directives, NEON intrinsics, and assembly code. In our case, the best decision is to use intrinsics. They provide access to all vector instructions and allows them to apply total control of instruction flow comparing with the auto-vectorization. The efficiency of such code is close to the assembly one.

For matrices processing, data parallelism may be used in two different ways. In the first case, the elements of the matrix are loaded into the vector register horizontally, first elements of the row with the indexes from 0 to 3, then from 4 to 7 and so on. For each subset of data vector instruction multiplication with accumulation is used. After the loop is finished, the dot product is obtained as a sum of four elements of the vector register. This method may be called horizontal vectorization and is simple for realization both for matrices multiplication and matrix inversion.

The second approach that may be called vertical vectorization requires source matrices to hold data in a vector format of type float32x4_t in columns (https://community.arm.com/processors/b/blog/posts/coding-for-neon---part-3-matrix-multiplication). In this case, code for matrices multiplication is very simple and efficient. But during matrix inversion, non-vector variables are used rather often.

To compare these variants of vectorization, the programs were made for both of them. These programs run on CubieBoard3 and Odroid-C2 computers with the same initial data as previous programs. Resulting plots of these tests are shown in Figure 7 and 8.

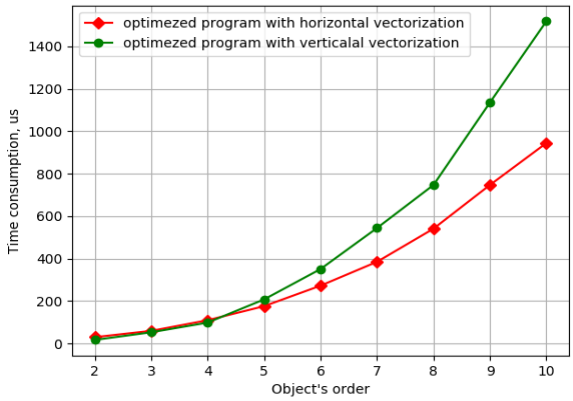


Figure 7. Time consumption for LSQ parameter estimation against object's order by the programs using vector unit NEON on CubieBoard-3.

As one can see in both cases horizontal vectorization gives better results. And further improvements and checks will be connected only with it.

3.7. Data Alignment

The next problem is data alignment. Old gcc compilers for armv7 required the directives explicitly showing that elements of the matrix rows and vector variables are aligned in memory at the boundaries multiple for 64 (http://infocenter.arm.com/help/index.jsp?topic=/com.arm.doc.dd

i0344k/Cihejdic.html). To find out the effect of the data alignment program with horizontal vectorization was modified. Compiler's directives __attribute__((aligned (64))) and builtin_assume aligned were addad to it.

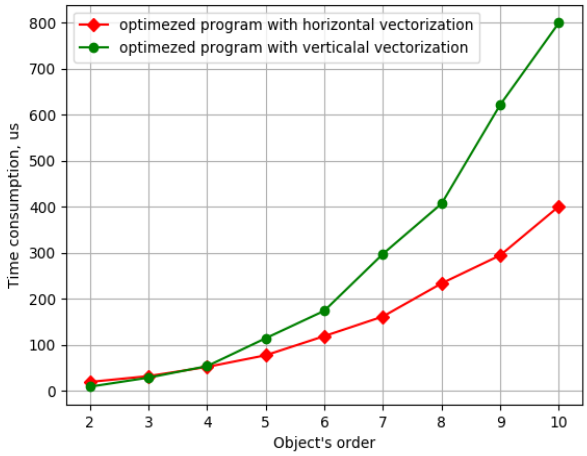


Figure 8. Time consumption for LSQ parameter estimation against object's order by the programs using vector unit NEON on Odroid-C2.

For comparison, together with 64, the numbers 8, 16, 32 and 128 were used for alignment. Running these programs along with the program without alignment directives has shown that there is no any difference between their time consumption. That means that modern gcc compiler makes data alignment without the additional directives.

3.8. Working with the Leftovers

The next problem one always meets while working with the vectorization is leftovers. ARM NEON vector register holds 4 single precision floating point numbers. The matrices sizes in real tasks are not multiple to 4, so the leftovers which can't be loaded into the vector register directly must be processed in some manner. Two ways to solve this problem are suggested (https://community.arm.com/developer/ip-products/processors/b/processors-ip-blog/posts/coding-for-neon---part-2-dealing-with-leftovers). The first is to process the leftovers as non-vector data. And the second is to extend matrices to the sizes multiple to 4 and fill the edges with the zeros. The second way is considered to be faster because the vector's processing is not interrupted with the non-vector operations.

In our case using extended matrices makes the code more sophisticated. Function for parameter evaluation takes two additional parameters and intermediate matrix S, holding informational and covariance square matrices side by side must be filled and processed in a not obvious way. As a result, the code of the function is tightly coupled with the rest of a program.

To avoid such a problem, the third way to solve leftovers problem were suggested. In this case function for parameter evaluation has local floating point arrays with the sizes equal to 4. The leftovers are loaded into these arrays before the main cycle of processing and the arrays are processed after the main cycle. These local vectors are extending each row of the matrix in a turn. As a result, all specific features connected with the vectorization are hidden within the function.

All three programs with different ways of solving the leftovers problem were checked for time consumption. The results are presented in Figures 9 and 10 for CubieBoard-3 and Odroid-C2 correspondingly.

Plots in Figures 9 and 10 show that variant with the local vectors is the worst from the point of view of time consumption, and the best is variant with the extended matrices. The difference for the object's order from 4 to 6 is about 50% for armv8 processor and 75% for armv7. This difference is significant, so using extended matrices for the solving of leftovers problem must be recommended.

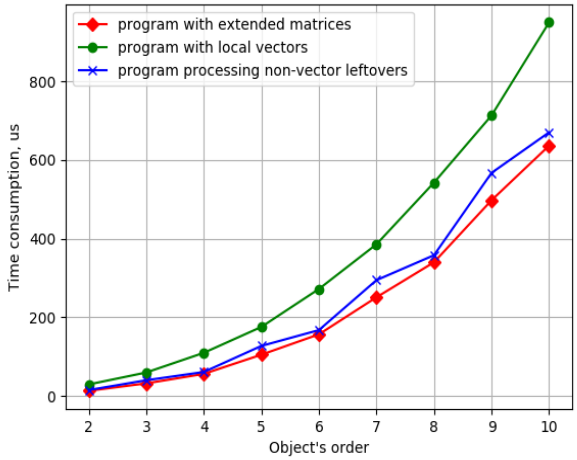


Figure 9. Time consumption for LSQ parameter estimation against object's order by the programs using vector unit NEON and different ways of leftovers handling on CubieBoard-3.

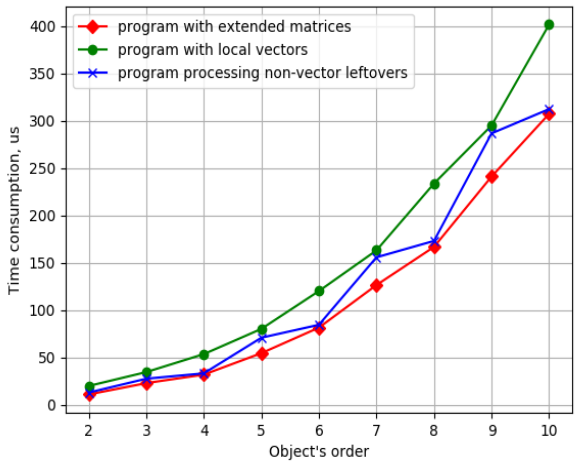


Figure 10. Time consumption for LSQ parameter estimation against object's order by the programs using vector unit NEON and different ways of leftovers handling on Odroid-C2.

3.9. Compare of all Results Obtained

All experimental data received in this work are presented in Table 2.

Types of the programs: 1 – QR-decompozition and dgels form lapack library; 2 – QR-decompozition and lsqsolve form gsl library; 3 – direct realization with the double precision numbers; 4 - direct realization with single precision matrices; 5

- direct realization with single precision matrices and multithreading; 6 – optimized function code; 7 – optimized function code and multithreading; 8 optimized function code with horizontal vectorization; 9 – optimized function code with vertical vectorization; 10 – optimized function code with horizontal vectorization and extended matrices; 11 – optimized function code with horizontal vectorization and non-vector leftovers.

Table 2. Time Consumption in μs for LSQ Parametrical Identification on Single-Board Computers CubieBoard-3 and Odroid-C2 for different Types of the Programs

Type	Object's order								
	2	3	4	5	6	7	8	9	10
CubieBoard-3									
1	675,8	675.0	765.4	897,4	1086,1	1250,7	1568,6	1873,9	2255,9
2	295.6	399.3	447.3	565.0	736.3	999.8	1399.4	1728.7	2228.7
3	64.7	198.3	243.2	452.7	765.3	1259.5	1915.6	2852.7	3925.8
4	37.5	92.3	191.4	355.8	596.5	981.2	1377.6	2125.9	2943.5
5	440.3	442.8	510.3	564.7	710.7	885.3	1270.5	1746.3	2100.6
6	18.8	50.5	110.5	215.2	348.6	545.3	828.2	1160.9	1551.4
7	419.0	440.4	461.9	483.3	592.3	718.2	915.9	1086.5	1329.3
8	29.5	59.8	109.5	175.7	272.3	384.7	540.6	746.7	944.2
9	16.5	52.7	98.3	207.5	350.6	543.8	746.6	1135.2	1518.2
10	13.02	31.8	56.0	104.8	156.5	250.7	339.8	497.4	636.3
11	14.9	39.8	60.8	127.1	167.3	294.3	357.7	567.0	670.1
Odroid-C2									
1	200.0	222.7	251.0	295.0	357.3	436.3	542.3	669.7	827.0
2	111.3	130.0	163.7	211.7	280.0	370.0	487.3	639.7	894.0
3	23.7	57.6	121.0	231.6	391.3	655.0	1031.7	1432.3	2007.3
4	24.0	57.0	120.0	226.7	383.0	595.7	884.3	1243.7	1771.0
5	543.3	547.0	627.7	819.3	925.3	1266.7	1478.3	1955.0	2422.3
6	17.0	40.3	84.3	162.3	269.0	417.0	616.7	882.4	1203.4
7	254.6	271.0	281.7	304.0	330.0	418.7	462.0	548.0	633.4
8	19.3	32.0	52.0	77.3	119.0	161.7	233.7	294.7	401.0
9	9.0	28.7	53.7	114.3	174.0	297.7	407.0	622.0	800.0
10	10.7	22.7	31.7	54.3	81.3	126.3	166.7	241.0	307.5
11	12.7	27.33	33.0	70.7	84.3	155.7	173.0	286.6	312.5

As Table 2 shows, the measures taken to optimize the code of function for LSQ parameter estimation of the object of control have allowed to reduce time consumption in 4.7 times for the processor armv7 and in 3.5 times for armv8 for the objects with the order from 4 to 6. Such significant productivity-increasing may allow to widen substantially the area of utilization of the adaptive digital controllers.

The sourse code of the programs tested in this article is available for free downloading under GPL license (https://github.com/basv0/lsq_armv7).

4. Conclusion

Wide using of the microprocessors systems that may provide high quality of technical objects control is restrained with the complexity of parameters estimation of these objects. As a result, the adaptive control may take more time than the technological process allows in hard real-time systems. This conclusion follows from the results of time consumption comparison for LSQ parameter estimation by the programs using function dgel from the linear algebra library lapack, function lsqsolve from scientific library gsl and functions for matrix multiplication and inversion. The tests show that direct realization of the matrix operations is more preferable for the tasks of not big dimensions and that using a single-precision floating-point variable instead of double precision ones does not decrease the calculations accuracy for the objects with the order less than 10.

Applying multi-threading showed that it gives productivity-increasing only for the objects with the order higher than 7 while in practice objects with the order between 2 and 6 are mainly met. Increasing productivity in the whole range from 2 to 10 may be achieved by the code refactoring and using intrinsic functions for the vector computations.

An optimized function using horizontal vectorization and extended matrices with the dimensions multiple by 4 has shown the best results. And this function is recommended for practical utilization despite the fact that matrix extension makes the code more sophisticated.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] V.V. Olonichev, B. A. Staroverov and M. A. Smirnov, "Optimization of Program for Run Time Parametrical Identification for ARM Cortex Processors," *2018 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, Moscow, Russia, 2018, pp. 1-5. doi: 10.1109/ICIEAM.2018.8728595.

[2] K. Astrom and B. Wittenmark, Computer Controlled Systems. Prentice-Hall, Inc., 1984.

[3] R. Izerman. Digital Control Sytems. Springer-Verlag, Berlin, Heidelberg, New York, 1981.

[4] L. Ljung, System Identification: Theory for the User. Prentice-Hall, 1987.

[5] Z. Tan, H. Zhang, J. Sun et al., "Research on Identification Process of Nonlinear System Based on An Improved Recursive Least Squares Algorithm", Proceedings of the 31st Chinese Control and Decision Conference, CCDC 2019 8832530, pp. 1673-1678.

[6] M. Arehpanahi, H R. Jamalifard, "Time-varying magnetic field analysis using an improved meshless method based on interpolating moving least squares", *Iet Science Measurement Technology*, vol. 12, no. 6, pp. 816-820, May. 2018.

[7] H. Li, J. Zhang, J. Zou, "Improving the bound on the restricted isometry property constant in multiple orthogonal least squares", *IET Signal Processing*, vol. 12, no. 5, pp. 666-671, Apr. 2018.

[8] V. Skala, "A new formulation for total Least Square Error method in d-dimensional space with mapping to a parametric line ICNAAM", *2015 AIP Conf. Proc.1738*, pp. 480106–1-480106–4, 2016.

[9] H. Leslie, Handbook of Linear Algebra. – CRC Press, 2013.

[10] C.F. Jeff Wu, Michael S. Hamada Experiments: Planning, Analysis, And Optimization. Wiley, New Jersey, 2009.

[11] B.A. Staroverov, V.V. Olonichev and M.A. Smirnov, "Optimal samplinig period definition for the object identification using least squares method", Vestnik IGEU, vol 1, pp 62-69, 2014. (article in Russian with an abstract in English)

[12] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, Numerical Recipes in C: The Art of Scientific Computing. Cambrige University Press, Cambridge, New York, Port Chester, Melbourne, Sydney, 1992.

[13] U. Drepper. What Every Prorgammer Shoud Know About Memory. https://www.akkadia.org/drepper/cpumemory.pdf.

[14] B. Stroustrup, The C++ Programming language. Addison-Wesley, 1997.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 35-43 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Design and Validation of a Meter Band Rate in OpenFlow and OpenDaylight for Optimizing QoS

Mohamed Saleh Al Breiki*, Suiping Zhou, Yuan Roger Luo

Department of Computer Science, Middlesex University, NW4 4BT, United Kingdom

ARTICLE INFO

Article history:
Received: 15 October, 2019
Accepted: 01 February, 2020
Online: 09 March, 2020

Keywords:
Cloud Computing
Meter Band Rate
Network Function Virtualization
OpenDaylight
OpenFlow
Quality of Service
Software Defined Networks

ABSTRACT

Technological developments in the Internet and communications have created a vastly complex and dynamic context with diverse heterogeneous networks and fast growth of mobile devices and multimedia. As the Internet becomes the primary mode of communication for many organisations there is requirement to enhance quality of service (QoS) from heterogeneous systems and networks. Traditional networks such as TETRA have become increasingly incapable of addressing the demand for media rich, bandwidth intensive traffic flows and applications. Mission-critical multimedia over new generation mobile networks face QoS constraints. This research explores a novel solution for quality of service performance for streaming mission-critical video data in OpenFlow SDN networks. A Meter Band Rate Evaluation (MBE) mechanism is advanced that improves the native QoS capability of OpenFlow and OpenDaylight. The MBE is a physical component added to the OpenFlow meter table to evaluate and dynamically adjust traffic rates and allows the traffic volume to be specified relative to other traffic in the network. Its design and development are presented, and the mechanism is verified through a simulated experiment in an SDN testbed. The results identified that QoS performance experienced a significant percentage increase when the MBE was active. These findings contribute a novel Meter Band Rate Evaluation mechanism that extends the native capability of OpenFlow and OpenDaylight to enhance the efficiency of QoS provision.

1. Introduction

Software Defined Networking (SDN) has received increasing attention as a new approach implementing new and innovative cloud-based architectures capable of integrating quality of service (QoS) enhancing end-to-end video performance. QoS dimensions of SDN and Network Function Virtualization (NFV) have the potential to address reliability and the quality of mission-critical applications over hybrid networks including 4G, 5G, Terrestrial Trunked Radio (TETRA) [1]. SDN is a novel technology designed to be dynamic and flexible to modern computing contexts by offering improved network management features [2] and enhanced network control for cloud computing [3].

The context of this research is UAE law enforcement in which there has been a growing dependence on mobile real-time video communications. TETRA networks have become increasingly incapable of addressing the demand for media rich, bandwidth intensive traffic flows and applications [4]. Modern policing like all areas of society is engaged in intensive sharing of different forms of real-time data and media across a diverse range of devices and platforms that form a critical part of its operations [5]. This context underlines the need to enlarge policing capabilities through the adoption of commercial mobile broadband and wireless networks and leveraging of cloud computing and SDN and NFV technologies to map multimedia applications over heterogeneous underlying networks. The programmability and the abstraction of the control plane from the forward and data plane is a key aspect of SDN's design [6,7]. This enables a single software control programme to control the data plane of numerous network elements [8] and enables greater and more flexible control and management of network components.

This paper contributes a novel Meter Band Rate Evaluation mechanism that extends the native capability of OpenFlow and OpenDaylight to enhance the efficiency of QoS provision. This research experiments with and analyses the implementation of an integrated set of technologies (SDN OpenDaylight controller, OpenFlow protocol and OpenQoS) over a centralized cloud platform using real-time video to 4G mobile devices. The primary goal of this research is to evaluate and validate the application of SDN and OpenFlow in realizing QoS for real-time video data transfer in a hybrid cloud computing context. A key objective is to

*Mohamed Saleh Al Breiki, +1 646 580 6284, ma2634@live.mdx.ac.uk

monitor QoS parameters and evaluate the SDN-based architecture in optimizing QoS performance. This paper presents a novel QoS mechanism that expands capabilities by incorporating a new meter band rate component. The experiment results in this paper verify the performance of this mechanism. This contributes a new mechanism and demonstrates the potential to expand the features of OpenDaylight SDN controller to realize more optimized QoS for real-time video traffic over heterogeneous networks. Thus this research represents a novel QoS mechanism that advances existing QoS capabilities of OpenFlow and OpenDaylight.

2. Related Work

Multiple studies have focused on improving quality of service (QoS) by means of leveraging the network control components of SDN by focusing software solutions at the switch level. This allows for transparency of network resource interfaces in alignment with application requirements.

Quality of service concerns the extent to which applications receive expected service quality relative to their level of mission-criticalness [1,9]. Hakiri and Berthou [1] indicate that QoS provision lacks this responsiveness to guarantee the required service level. The design of the OpenFlow protocol allows the SDN controller to access and modify flow data at the switch level for all devices thereby connecting the control and data planes [10]. OpenFlow flexibility allows for data to be automatically directed with minimal disruption achieved by decoupling the forwarding layer from the control layers [11] which manages data flows and creates or updates rules for devices in the network [12].

Changes in hardware are not required by employing OpenFlow and vendors do not need to open systems to support it provided a small number of widely used interfaces are available to the control plane [13]. Any system that integrates the protocol is capable of controlling the switch behaviour of any devices in the Openflow network. Directives allow ISPs to deploy SDN to guarantee quality of service of streaming services alongside traditional network services [13].

The significant potential to manage QoS within the OpenFlow protocol has generated multiple studies focused on addressing QoS issues at network level. A key attribute of OpenFlow is the focus of QoS to address specifically different parameters such as delay, throughput and jitter [14] as a result of the control of packet and flow level activity. This granularity permits users of OpenFlow to specify how individual flows are to be managed in accordance with Integrated Services (IntServ) and open standards definitions [15]. Individual flows can also be amalgamated into classes by users. OpenFlow grants access to a set of programming tools which enable the generation and recycling of virtual flows or slices. Thus the user can stipulate how network resources including routers, switches and queues are allocated to different slices with different priorities [15].

In [2], the author proposes an SDN-based solution to ensure network-wide QoS based on multi-paths routing that utilizes OpenFlow's queue support to maximize the efficiency of resources in line with the network user requirements and minimizing excess provisioning. The solution focuses guaranteeing end-to-end QoS for individual data flows and the efficient management of open virtual switches (OVSs). The OVS configuration is recalculated in response to new network demands or as users progress on to new nodes in order to minimize the utilization of OVSs in the network.

In an attempt to guarantee bandwidth end-to-end Tomovic et al. [16] advance an explicit QoS guarantee as in IntServ by incorporating controls for admission and reserving bandwidth in OpenFlow. This commences by transmitting the bandwidth requirements of a QoS flow initiating a Constrained Shortest Path computation by the controller in which the required bandwidth is used as a limiter. The technique uses the available unreserved bandwidth as the link weight ensuring that the reserved flows are allocated to bandwidth that has been reserved. The flow is rejected where the bandwidth is not available for end-end path. Furthermore, there is intermittent checking by the controller of all links associated with the QoS flow, with utilization greater than 80% resulting in the re-routing of best-effort flow to an alternative path with lower utilization.

To maintain QoS in real-time video streaming in SDN Liang and Shen [17] introduce a new SDN-based routing approach that is capable of dynamically routing deployment across layers of Scalable Video Coding (SVC) accounting for bandwidth and different QoS parameters: delay, and packet loss. SDN research by Celenlioglu and Mantar [18] advanced a routing and resource control approach for enhancing QoS that takes into account pre-defined paths and reservation of resources between network segments. By improving routing scalability and decreasing admission time the solution provides QoS guarantees to stationary nodes.

Research into Openflow by Egilmez et al. [11] contributed an OpenQos controlled solution that concentrated on achieving end-to-end QoS which secures an optimal path for the service delivery of application traffic that prioritize multimedia traffic [19]. Flows for data, video and voice are categorized and subject to different algorithms either shortest routing for data or dynamic routing to guarantee QoS.

In a new approach [8] multiple issues are identified: the requirement for each node device to configure the same information; need for manual configuration of all ports of each forwarding device which can be error-prone and time-consuming; and the incapacity of legacy networks to make a distinction between the different resource requirements of voice and data network traffic. Data traffic requires more bandwidth while voice is associated with small packets and delay sensitivity. Achieving QoS in SDN is predicated on two key steps of marking packets for the data plane and management at the control plane. The network is made up of multiple virtual networks whose resources and requirements are autonomously monitored and managed [8].

To enhance QoS some research has focused on an extension of OpenFlow. The purpose of research by Lu et al. [20] was to provide a solution for QoS using conventional flow control techniques and maximizing flow. A QoS slice method and forwarding mechanism is advanced nevertheless the study does not address the issue of supporting QoS in OpenFlow natively. Other studies undertaken in this field do not examine the constraints associated with QoS by OpenFlow [8, 21, 22, 23]. Bhattacharya and Das [24] propose that a module is embedded to minimize any communication latency between switch and controller by means of caching flows nevertheless a QoS mechanism for the data plane is not advanced.

Examining the utilization of DiffServ and multipath QoS routing in OpenFlow Jinyao et al. [25] propose HiQoS as a method to guarantee QoS. In HiQoS traffic is classified into three groups of best effort data stream, low delay interactive multimedia and high throughput video streaming. The switches forward the different categories of traffic into three preconfigured queues with fixed rates that effectively divides bandwidth. Flow paths are determined by employing existing bandwidth usage as weight and flows from the same category of traffic are able to utilize different paths. While the method distinguishes low delay traffic as a distinct category it fails to account for delays in relay-time and calculates routes using existing bandwidth information. The assumption is that greater utilization in terms of bandwidth and queue leads to greater delay. Therefore HiQoS aims to guarantee bandwidth for categories of traffic rather than providing guarantees at flow level [25].

3. Problem Formulation

While cloud networking solutions based on SDN have specific performance, scalability and flexibility advantages, they are nevertheless limited to supporting specific kinds of network environments and lack more precise control of network paths for attaining objectives such as traffic prioritization or rapid failover [11]. QoS mechanisms including Differentiated Services (DiffServ), IntServ and Multiprotocol Label Switching (MPLS), are subject to numerous limitations in respect of level of flexibility and integration in different systems. The potential of SDN is associated with the ability to extend the diversity and targeting of resource provision so that it better aligns with specific application of QoS requirements [26]. One component of this is a suite of tools that facilitate distributed and collaborative management to achieve better classification of data flows in accordance with traffic characteristics.

The OpenDaylight SDN controller based on the OpenFlow protocol is widely used and has been embraced by vendors worldwide. OpenFlow is a key component within SDN as it controls switches through signalling between switches, guarantees the provision of QoS and is significant in respect of the best effort paradigm. The major benefit of the protocol is backwards compatibility and its ability to be integrated within current network infrastructures and legacy flow-enabled network devices with minimal changes in design. Outlining the present capabilities of OpenFlow's support for QoS in SDN assists understanding of the problem. Currently support is provided using two approaches of either DiffServ or per-flow QoS. Using Per-port Queue, a queue for a flow in a flow-entry can be specified within OpenFlow so that in the event of a match decision is allocated for each queue. Every single queue is configurable with properties, governing the treatment of packets within the queue. OpenFlow's per-flow queue attribute is utilized for the definition of per-flow QoS by means of a per-flow meter table in which single or multiple flows can be attached. The role of meters is to determine packet rate and control that rate. There is also the possibility to attach single flows to multiple meter tables on the condition that different tables are utilized. Within a single table each flow may be assigned to only one meter table. The combination of per-flow queues and meter tables can enable the implementation of DiffServ.

There are a number of drawbacks and areas in which QoS in OpenFlow (versions 1.1-1.5.1.) could be improved. Simple queuing mechanisms such as min-rate and max-rate provide limited support for QoS and queues cannot be configured inside OpenFlow. Only rudimentary support for DiffServ (for example, DSCP Remark) is available. Moreover when using queue, QoS is solely permitted at egress. Lastly, most switches do not support optional meters. The simplicity of the queueing mechanisms engenders two key problems for QoS. Firstly the per-port queue allows the available bandwidth to be sliced into a number of queues and to specify certain properties for each queue of min-rate, max-rate and others. However this action fails to guarantee that the bandwidth available is efficiently utilized. Furthermore over-provisioning to avoid low QoS during congestion may be needed, resulting from the specification of minimum bandwidth for traffic. Behavior is reliant on the particular hardware implemented and is therefore not able to be predicted. A second QoS issue is the restricted scalability entailed by the simple queuing mechanisms. For all traffic, there is a flow-entry requirement to specify which it belongs to however switches have a limitation in the number of flow-entries. For a large network, every switch needs flow-entry and queue configuration for all traffic in the network. Furthermore, OpenFlow's inability to allow for queues to be configured implies that dynamic changes cannot be made. This results from the nature of the SDN controller which is not able to be utilized to configure queues and therefore other mechanisms including OVSDB, OF-Config, or cli must be deployed. These however do not work in the fast path. Consequently end-to-end QoS implementation is challenging using this approach. The basic support provided for DiffServ such as DSCP Remark is a further key issue undermining QoS. Fields are adjustable by means of meter tables nevertheless meters complement rather than replace queues and are unable to surmount the limitations associated with them. Within a packet meters have the capacity to alter certain fields but do not provide support for the comprehensive operation of all kinds of per-hop-behavior (PHB). Illustrating this point, different queuing mechanisms such as strict priority queuing (SPQ) or weighted fair queueing (WFQ) are not able to be applied as solutions for a PHB. In addition there are feasibility issues in applying mechanisms such as Call Admission Control (CAC) when using meters. By permitting QoS at the single point of egress only the use of queues is problematic as there is a distinct advantage in the ability to drop traffic at the point of ingress that safeguards switch resources, when possible, for other traffic. Meters can support this however possess a limited capability for QoS as previously outlined. The prevailing absence of switch support for meters underlines a lack of guarantee that support will be provided by OpenFlow switches.

4. Experimental Study

This research applied an experimental approach in order to test and analyse video streaming quality of service performance under different experimental conditions. This comprised the establishment of an SDN based network testbed based on OpenDaylight SDN controller and OpenFlow 1.3 and utilizing the Mininet emulator to undertake several tests. The testbed was implemented in line with the system parameter and mechanisms indicated and detailed in Figure 1 and Table 1 that specify the key software versions utilized during the implementation.

Table 1: Software versions utilized in implementation

Software	Description
Python 2.7.14	Programming language
Linux	Host OS
Open vSwitch 2.5	Virtual multilayer switch with OpenFlow support
Ubuntu 16.04	Virtual machines
Yang UI	User interface for OpenDaylight
Mobile Phone Samsung 4/4G	Client smartphone device
Networks	WiFi (WLAN), 3G, 4G, SDN Controlled
OpenDaylight Controller	SDN framework
Mininet 2.2.1	Network topology environment

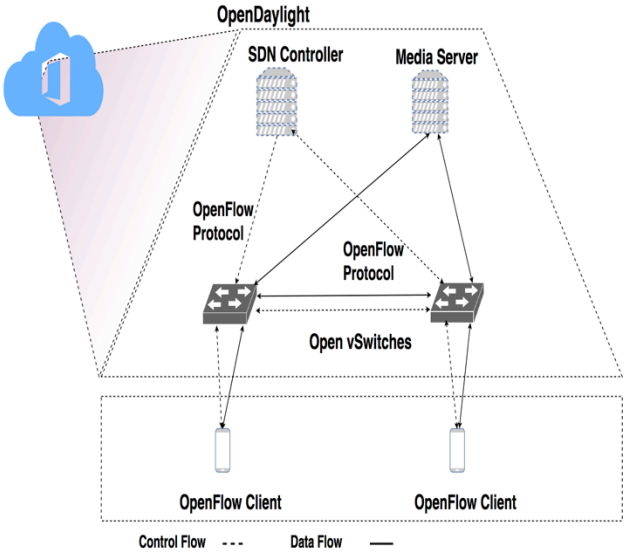


Figure 1: OpenFlow-based SDN network testbed

OpenDaylight Controller (ODL) was utilized as the implementation framework subsequent to assessment of different controllers. Key benefits are associated with ODL in terms of: regular updating relative to other controllers; support for all versions of OpenFlow; necessary features to allow for the implementation of a traffic engineering platform in SDN, and possession of several existing modules able to be used with customized applications. Firstly a series of experiments were undertaken to assess QoS performance for existing SDN/OpenFlow capabilities in a high traffic environment utilizing two distinct sets of experimental conditions. The experiment allowed for detection and analysis of issues associated with QoS and the constraints of SDN/OpenFlow that can be addressed in the following stage of research. The tests provided benchmarks for different QoS metrics of delay, packet delay variance (jitter), packet loss and throughput.

5. Meter Band Rate Design

The design of the Meter Band Rate algorithm is focused on enhancing the native capability of OpenFlow at the switch level. The design of the proposed mechanism is outlined in this section and an overview of MBR model and components is represented in Figure 2. The evaluator represents the novel component of this design based on a band rate description language (BRDL) for analysing required dynamic states for the mechanism. The BRDL facilitates classification of network traffic. For instance, if there are i = 1 . . c types of classes in a network with priority of types as follows, 1>2>…>c such that the network is sliced into c classes. The BRDL syntax enables the specification of slices relative to others within the network:

$$\text{rate}(x) = \{\text{MAX-}\}\{\text{rate}(y)\text{-}\} \; . \; . \; \{\text{rate_const}\} \quad (1)$$

Where, MAX is the total available capacity, rate(x) specifies rate of this traffic class, rate(y) specifies rate of any other traffic class and rate_const is any constant. Bandwidth reservation is controlled by rate_const and avoids the need for explicit assignments of services to meters. In the case of different types of traffic, for instance voice or data flows, the data traffic needs only to be configured with meter and assigned the necessary bandwidth for voice traffic to rate_const. To exemplify the value of this consider two types of flows in the network of voice traffic and iSCSI traffic. Voice flow has a fixed rate and therefore can be assigned a constant rate. Therefore, only rate_const part will be applied:

$$\text{rate(voice)} = \text{rate_const_voice} \quad (2)$$

Then, iSCSI traffic is assigned rate(y)=rate(voice) is set and the rate_const component is not required. Therefore,

$$\text{rate(iSCSI)} = \text{MAX} - \text{rate(voice)} \quad (3)$$

Using this approach, the MBR invocation is either DROP or ALLOW. The benefit of the BRDL is in overcoming the limitation of meters and per-port queues by adding constant max-rate and min-rate limiting features. Further this optimizes traffic flow in terms of throughput and simultaneously guaranteeing QoS requirements of each individual slice. The SDN controller assigns all flows with OpenFlow meter at the install stage with unique meter_id identifier associated with local switch for each flow entry. This action is dependent on the supported meter band type:

1. OFPMBT_DROP
2. OFPMBT_DSCP_REMARK
3. OFPMBT_EXPERIMENTER

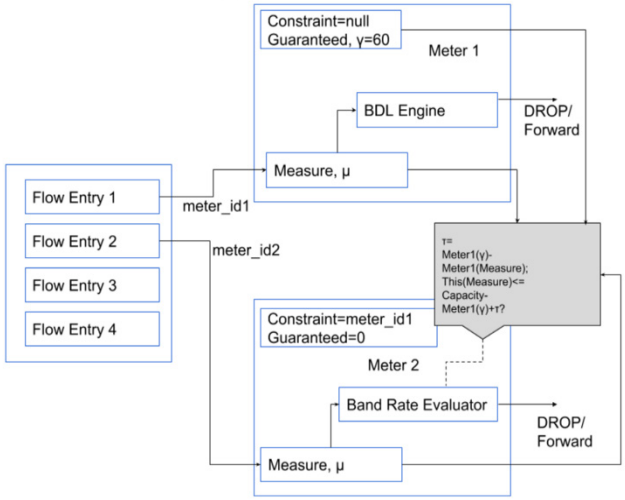


Figure 2: Meter Band Rate model

The OFPMBT_EXPERIMENTER indicates the band rate evaluator in this model. OpenFlow meter bands comprise three bands: type, rate, burst and counters. *Constraint* and *guaranteed rate* extends the meter band to enforce priorities. The *constraint* records the priority of meter_id's supplied by band description language given in Equation (1). Thus the *constraint* list of "band rate evaluator" is determined from *rate(y)* and *rate_const* of BRDL. The *guaranteed rate* specifies the bandwidth guarantee for this meter. The *measure* field reflects the measured data rate by the meter. The above values can be specified in KBPS or packets per second as set in meter flags.

Figure 2 depicts the flow-entries associated with meters. Flow 1 is associated with Meter 1 having meter id "meter_id1" and Flow 2 is associated with Meter 2 having meter id "meter_id2". The *band rate evaluator logic* is given in the following section. Thus all incoming packets are associated with a meter flow entry at the ingress port. Figure 3 depicts meter evaluation to determine the band rate to drop the packet, so that where the evaluated rate is higher than the existing rate, the packet is passed.

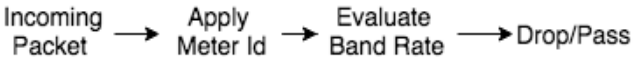


Figure 3: Applying Meter Band Rate Evaluator to all packets

Figure 4 illustrates the evaluation process that treats guaranteed and measure components as *superior constraints*. The evaluator meter uses the *guaranteed* and *measure* parameters of the meters specified in the *constraint* field. Parameters of "higher priority" or "superior" meters, are considered *superior constraints*. The constraint meter list is used to identify superior constraints. Step 1 - the measure field is subtracted from the guaranteed field of each meter in the constraint list, which yields the available bandwidth of each constraint meter; Step 2 - available bandwidths are totalled; Step 3 - the band rate evaluator adds the guaranteed fields and deducts the result from the total port capacity; Step 4 - the results from step 2 and step 3 are totalled. The result from the step 4 is compared with the measure field; if higher, the packet is allowed to pass. In respect of compilation of the band description language into the constraint list used by band rate evaluator three types of traffic are specified: voice, video and data. The data rates of these traffic types in band description language can be expressed as:

rate(voice) = 32Kbps (const)
rate(video) = 1Mbps
rate(data) = MAX – rate(video) – 32Kbps

The rate_const (=32Kbps) is used in expressing rate (data). Meters will not be required for voice traffic. Thus, there are two meters: meter_video and meter_data. The Constraint and Guaranteed fields of these meters are as follows:

meter_video: {constraint = null, guaranteed = 1Mbps}
meter_data: {constraint = meter_video, 32Kbps}

The proposed design focuses on the limitations of OpenFlow to provide a flexible and optimum QoS solution. Currently there are two instruments inside OpenFlow to harness traffic behavior, namely per-port-queues and meters. Both queues and meters have their own limitations in terms of versatility and efficiency.

To address these shortcomings the following components have been incorporated in the proposed solution:

1) Band description language (BDL)

2) Band rate evaluator (BRE)

3) Constraint list

4) Guaranteed rate

The band description language (BDL) is a logical component of the design which allows the network administrator to specify desired QoS behavior in a flexible and efficient manner. A major limitation of OpenFlow queues is that they only allow for simple QoS behavior such as max-rate limit. This means that rate limited traffic will face restriction despite the availability of free capacity in the network. In this solution, BDL allows the traffic volume to be specified relative to other traffic in the network. As a result, there is no unnecessary bandwidth restriction for any traffic.

The inability to be provisioned in the fast path is another limitation of the queue-based approach. Consequently, the SDN controller is unable to update the queues in real time to reflect network demand. The proposed design solves this issue firstly by making BDL capable of specifying per-flow bandwidth in terms of other flows. Therefore, flow rates are automatically adjusted to network demand. Secondly, OpenFlow meters are used to implement the solution. Meters in OpenFlow can be configured in the fast path, therefore the QoS settings can be updated in real time by the SDN controller in the solution.

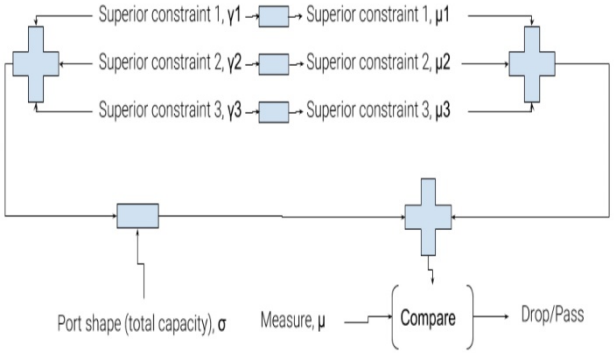


Figure 4: Evaluator process

The band rate evaluator (BRE) is a physical component added to the OpenFlow meter table. Presently, in OpenFlow, band rates can be specified as static constants. When the traffic rate reaches one of these band rates, OpenFlow meter will execute a certain action (e.g. dropping the packet, or modifying certain protocol header fields in the packet). This static meter bands feature can be useful for simple rate limiting but is unable to adjust to dynamic network demand. It is theoretically possible to periodically update these static band rates by the SDN controller. However, in practice the traffic rate is too fast for the SDN controller to be able to continuously update meter band rates, even in a small network. In

this solution therefore BRE is used to dynamically adjust traffic rates based on the BDL.

The constraint list is another physical component of the proposed design. As discussed, the BDL and BRE allow traffic rates to be specified relative to other flows. The constraint list is responsible for storing information in relation to which traffic rates are dependent upon which other traffic. The constraint list is designed to be understood by the BRE whereas the BDL is intended for use by humans and the SDN controller. The QoS rules specified in BDL are processed by the SDN controller to create the constraint list. Once the constraint list is generated, the SDN controller will install it into the SDN switch.

The guaranteed rate is the final physical component of the proposed solution. This allows reservation of a virtually dedicated amount of traffic for a flow. Use of the term 'virtually dedicated' signifies that this operation is not a dumb reservation as performed by min-rate limiters in OpenFlow queues. In OpenFlow queues, reserved bandwidth by a min-rate limiter will be unavailable to other traffic when the flow for which bandwidth is reserved has a lower rate than reserved bandwidth. This means that the reserved bandwidth will be wasted most of the time. In this design the BRE will mitigate this problem by allowing the bandwidth specified by the guaranteed rate for a certain flow to be re-used by other traffic as long as there is some free bandwidth unused by the owner of the guaranteed bandwidth. The guaranteed rate can be left unspecified for any flow as long as it is not mentioned in the constraint list of another flow. This can be useful for best effort traffic which requires least priority. For other traffic which has more priority compared to other traffic the guaranteed rate must be given.

6. Meter Band Rate Evaluator Implementation

The Meter Band Rate Evaluator outlined earlier was deployed in OpenFlow Software Switch 1.3. The module ofsoftswitch13 was modified to integrate the mechanism to implement the Meter Band Rate Evaluator. Table 2 depicts the key components and core functions of the Meter Band Rate Evaluator algorithm developed for this implementation. Several modifications were made to the OFLIB. All messages in OpenFlow share common header containing information message type where the default sruct for each message allows for two actions: marshalling (packing) and unmarshalling (unpacking). These functions enable Ofl-messages transmission of new modifications through wire format. In Ofl-struct.h the meter_configuration was modified with statistical function adding to the number of fields to perform metering monitoring of parent free tokens. A flag in meter configuration was added to facilitate enabling or disabling of the band evaluation engine. Oflib-messages were modified to allow for band evaluation engine to be set to enabled/disabled states. The central part of the algorithm is deployed in meter_entry.c of the udatapath component to enable parent monitoring support and to apply entries. The Meter Table is responsible for QoS operations and was adapted with fields to define new actions.

The meter table performs the main processing of packets in accordance with the band meter design. Meter_table.c was modified to activate the evaluation engine and increase the structure size from 16 to 20. Finally, the BEEFLAGS within the utilities/dpctl component was modified to enable tracking and debugging and query of Flow Table state in OpenFlow switches which avoids the requirement for debugging at the controller level.

Table 2: MBE core algorithm

Packet Handling at Meter

Input: *packet, monitoringPeriod*
Output: *decision*
 	
WHILE *parent* ← *nextParent* DO
IF *free(parent)* < *free*
free ← *free* + *free(parent)*
ENDIF
ENDWHILE
 	
elapsedTime ← currentTime-startTime
 	
IF *elapsedTime* = *monitoringPeriod*
totalFree ← free
startTime ← *currentTime*
ENDIF
 	
IF *available(meter)* > *size(packet)*
decision ← *ALLOW*
available(meter) ← *available(meter)-size(packet)*
ELSIF *free* > *size(packet)*
decision ← *ALLOW*
free ← *free-size(packet)*
ELSE
decision ← *DROP*
ENDIF

7. Results

Two rounds of testing were undertaken with and without the Meter Band Rate Evaluator (MBE). Initial tests conducted without the MBE provided a benchmark for performance and gathered and analysed QoS data for four parameters of delay, jitter, packet loss and throughput under conditions simulating a congested environment. The second stage of testing was undertaken after the MBE had been installed and configured in order to evaluate the capacity of the novel solution to improve QoS performance. The tests undertaken in the first experiment were replicated under the same environmental conditions for periods of 4.3 minutes.

For the four QoS parameters the results showed that there was a marked improvement when the MBE was applied. The figures chart the comparison between the benchmark tests and those for the MBE. Figure 5 reveals significant optimization in delay performance with a reduction in delay time over the period of 260 seconds. In the first round of tests delay ranged on average between 500-1100ms in comparison to results with the MBE enabled of between 80-360ms. Average delay with the MBE disabled was 546.7ms compared to 206.6ms with the MBE.

For the remaining three QoS performance measures the positive impact of the MBE was evident and provides support for the solution to optimize QoS performance within OpenFlow. Analysis of the data for jitter performance as shown in Figure 6

indicates that in comparison application of the MBE significantly reduced jitter to a median average of 1430 compared to 1906.5 without the MBE. Similarly there was a substantial minimization of packet loss when the MBE was active as shown in Figure 7, indicating a consistent average of under 5% across all tests, compared to between 5-15% when the MBE was inactive. The median average PLR was 2.1% when the MBE was enabled compared to a figure of 8.3% when disabled.

Results also revealed significant optimization of throughput performance, a key QoS performance measure for ensuring the quality of video. As shown in Figure 8 when the MBE is active there is a significant increase in throughput of more than 4000 packets/sec and rising to 9000 packets/sec, a markedly higher figure than when the MBE is inactive. Results without the MBE show that packets/sec varied between 300-4000 packets/sec.

As Table 3 shows for all QoS parameters there was a significant percentage change in performance when the MBE was active. For delay, results reveal an average reduction of 94% reflecting the substantial impact of MBE on performance for this measure. A mean percentage reduction in excess of 800% for jitter is indicated, and when extreme values are accounted for a 32% median average decrease in jitter is shown.

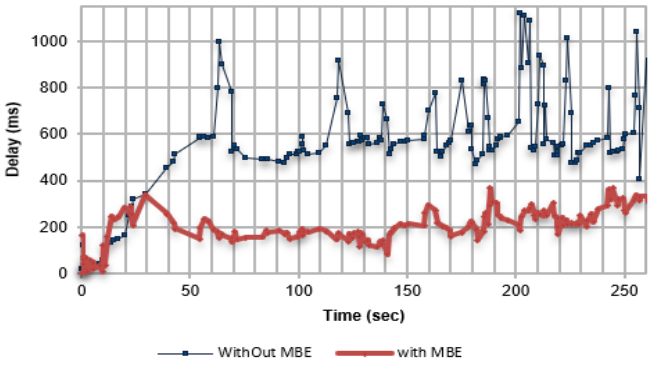


Figure 5: Delay performance with MBE

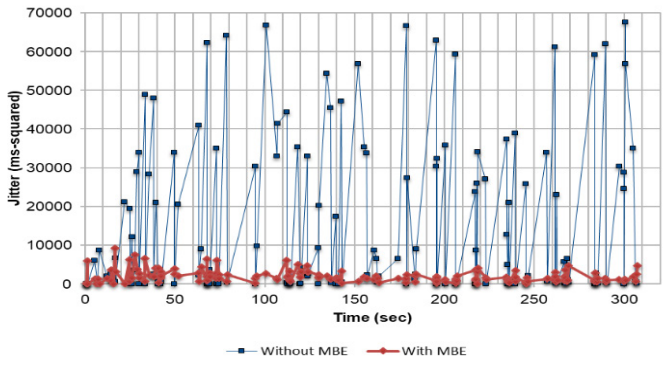


Figure 6: Jitter performance with MBE

Table 3: Percentage change in QoS measures with Meter Band Rate Evaluator

	Delay	Jitter	Packet Loss	Throughout
Mean % Change	-94.16	-866.45	-285.32	200.20
Median % Change	-96.27	-32.76	-325.83	189.49

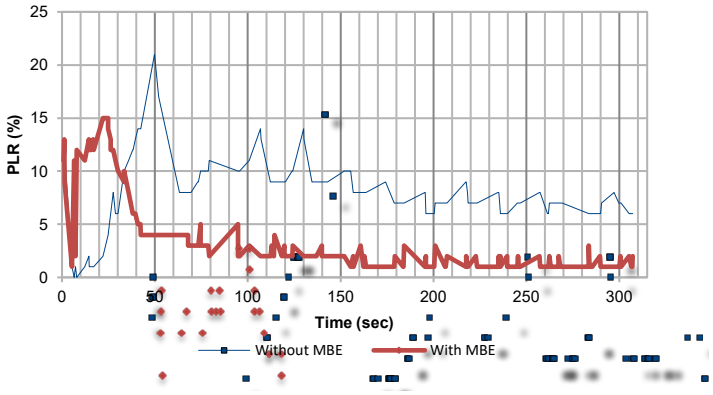


Figure 7: Packet loss ratio performance with MBE

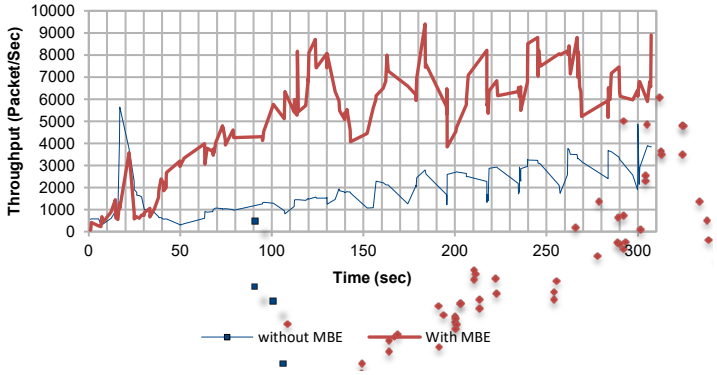


Figure 8: Throughput with MBE

In relation to ratio of packet loss the mean average decrease was 285% and median average decrease of 325%. Lastly, throughput under MBE increased by 200% for the mean average and 189% for the median. Based on the metrics tested these results provide substantial validation and support for the ability of the MBE to improve the native capability of OpenFlow QoS performance.

8. Discussion of Findings

This research explored application of an SDN-based mechanism to enhance QoS for real-time end-to-end video transmission. The solution centres on the development of a Meter Band Rate Evaluator (MBE) mechanism employing a new band rate description language that extends the capabilities of OpenFlow and OpenDaylight to support QoS. This solution was designed and subject to experiments in a Mininet test environment which revealed significant improvements in level of QoS performance across multiple parameters: jitter, throughput, delay, packet loss. This contribution addresses the limitations of existing methodologies to guarantee QoS performance, either through marking data-plane packets and management at the control plane automatic manager for monitoring requirements of virtual networks [8] or slice and forward methods [20]. The findings of this study demonstrate potential to address the limitations of QoS in OpenFlow's native features and protocol to support QoS [8, 21, 22, 23]. Bhattacharya and Das [24] introduce an embedded module to reduce communication latency between switch and controller by caching flows, however a data plane QoS mechanism is not proposed.

The mechanism applies a hierarchical approach in classifying traffic classes providing higher positions of priority for mission-critical types of data. This provides flexibility for defining relative QoS relationship between different classes of traffic. The meter band rate determines the point of application using constant 32-bit values specified either in packets/sec or kb/s which can be updated by OFPT_METER_MOD message. This constant inhibits the application of a more dynamic and QoS responsive approach because it fails to account for the co-existence of flows within the wider networks. This solution addresses this issue by specifying band rates of slices relative to multiple slices in the wider network. The issue of constant rate limiting in existing meters and per-port queues is addressed. Furthermore, the results validate performance of this new mechanism to provide higher levels of QoS across the four parameters measured.

In comparison with other solutions in the literature this design offers some distinct contributions. Civanlar et al. [27] aim to provide QoS guarantee to video services similar to the proposed architecture in this solution, however their work is limited to selecting best route for QoS traffic. Minimal detail is provided on how to prioritize the QoS traffic in the SDN infrastructure layer, except for the assumption that QoS traffic can have the right to disrupt non-QoS traffic when they are placed on the same route. The primary focus of this work is on overall best path selection therefore how traffic is processed at each hop is not considered. Contrastingly the proposed solution in this paper focuses on this area and provides per-hop processing with the help of the four physical components.

In [28] video traffic is divided into two classes (lossless and lossy) accompanied by two types of routing algorithms generating different types of flows. The number of traffic classes is shown to have significant impact on overall video performance. From this it can be concluded that there is an advantage to flexibility in the number of classes that can be implemented. The proposed algorithm is flexible enough to accommodate any number of traffic classes as in BDL there is no limitation to the number of meters. Therefore, the network administrator is free to design the system with his or her preferred number of classes.

In advancing OpenQoS Egilmez et al. [29] extend the OpenFlow controller to provide QoS for multimedia traffic. In this architecture the call admission process is relied on to block incoming traffic when there is insufficient resource available in the network to guarantee the requested QoS. However, this could deny admission to voice traffic although possible to guarantee QoS for this type of traffic at the expense of quality for non-priority traffic. The protocol proposed here does not suffer from this limitation of OpenQoS as there is no call admission process. QoS is guaranteed by using OpenFlow meters and the band evaluation engine.

Certain limitations are acknowledged in relation to the proposed solution. The maintenance of network state information is increasingly complex under the dynamic character of the network typology and changing route behaviour. Established routing paths can be disrupted even while data is being transferred entailing, for QoS, necessary maintenance and development of paths in ad hoc networks that have minimal routing overhead and delay [30]. Mobile nodes may further be constrained by limited power supply in comparison with the nodes of wired networks. Ensuring QoS could mean that, as a result of overheads, more power is used that rapidly drains that available to the node [30]. Scalability of this solution may also be impacted by the computational cost associated with the centralized algorithm in SDN. As networks become larger in size, they become more expensive and difficult to maintain requiring greater storage and more powerful infrastructure [31]. Finally increased computational time on systems in the network can impact negatively on the energy consumption of user devices [31].

9. Conclusion

These findings contribute a novel Meter Band Rate Evaluation mechanism that extends the native capability of OpenFlow and OpenDaylight to enhance the efficiency of QoS provision. The focus of this research was on addressing constant rate limiting with existing QoS mechanisms by extending the capability of OpenFlow to support efficient QoS provision. The introduction of a Meter Band Rate mechanism extends knowledge and suggests new possibilities in developing the existing capabilities of OpenFlow and OpenDaylight to support QoS. Experiment results indicated a significant increase in video streaming and multimedia traffic for multiple QoS parameters and point to the scope of the meter band rate mechanism for optimization of QoS in OpenFlow. The evaluator represents the novel component of the design based on a band rate description language (BRDL) for analysing required dynamic states. This allows specification of desired QoS behavior in a flexible and efficient manner and there is no unnecessary bandwidth restriction for any traffic. Further research in this area may subject the mechanisms to further testing and validation across more complex networking scenarios and hybrid networks. In addition, this may address a comprehensive range of optimization strategies to establish performance and validation in relation to load balancing, energy-consumption, controllability of network flow, safe updating or network security.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The research has been supported and sponsored by the Ministry of Interior in the United Arab Emirates as part of an international research programme.

References

[1] A. Hakiri, and P. Berthou, “Leveraging SDN for the 5G networks: Trends, prospects and challenges,” 2015, http://arxiv.org/abs/1506.02876

[2] D.L.C Dutra, M. Bagaa, T. Taleb, and K. Samdanis, “Ensuring end-to-end QoS based on multi-paths routing using SDN technology,” in GLOBECOM IEEE Global Comms Conf., 2017, pp. 1-6.

[3] M. Banikazemi, D. Olshefski, A. Shaikh, J. Tracey, and G. Wang, “Meridian: An SDN platform for cloud network services,” IEEE Comm. Magazine, vol 51, pp. 120-127, 2013.

[4] J. Dunlop, D. Girma, and J. Irvine, Digital Mobile Communications and the TETRA System, John Wiley & Sons, 2013.

[5] R. Blom, P. de Bruin, J. Eman, M. Folke, H. Hannu, M. Näslund, M. Stålnacke, and P. Synnergren, “Public safety communication using commercial cellular technology.” in IEEE Next Gen Mobile Applications, Services and Technologies, NGMAST'08., 2008, pp. 291-296.

[6] M. Jammal, T. Singh, A. Shami, R. Asal,. and Y. Li, “Software defined networking: State of the art and research challenges,” Computer Networks, vol 72, pp.74-98, 2014.

[7] O. Coker, and S. Azodolmolky, Software-Defined Networking with OpenFlow, Packt Publishing, 2017.
[8] J. Shi, and S.H. Chung, "A traffic-aware quality-of-service control mechanism for software-defined networking-based virtualized networks," Int. J. Dist. Sensor Networks, vol 13, pp.1-13, 2017.
[9] J. Huang, P.J. Wan, and D.Z. Du, "Criticality-and QoS-based multi-resource negotiation and adaptation," Real-Time Systems, vol 15, pp. 249-273, 1998.
[10] T. Reza, Learning OpenDaylight: A Gateway to SDN (Software Defined Networking) and NFV (Network Functions Virtualization) Ecosystem, Packt Publishing Ltd, 2017.
[11] H.E. Egilmez, S. Civanlar, and A.M Tekalp, "An optimization framework for QoS-enabled adaptive video streaming over OpenFlow networks," IEEE Trans.on Multimedia, vol 15, pp. 710–715, 2013.
[12] M. Sisov, "Building a software-defined networking system with OpenDaylight controller", BSc Thesis, Inf. Techn., Helsinki Metro. Uni. of Applied Sciences, 2016.
[13] I. Trajkovska, P. Aeschlimann, C. Marti, T.M. Bohnert, and J. Salvachúa, "SDN enabled QoS provision for online streaming services in residential ISP networks," in IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), pp. 33-34, 2014.
[14] S. Jivorasetkul, M. Shimamura, and K. Iida, "End-to-end header compression over software-defined networks: A low latency network architecture," in 4th Int Conf on Intell. Networking and Collab. Sys., (INCoS), pp. 493–494, 2012.
[15] F. Hu, Q. Hao, and K. Bao, K, "A survey on software-defined network and openflow: From concept to implementation," IEEE Comm. Sur. & Tutor., vol 16, pp. 2181-2206, 2014.
[16] S. Tomovic, N. Prasad, and I. Radusinovic. "SDN control framework for QoS provisioning,". Telecomms Forum Telfor (TELFOR), 22nd. IEEE. pp. 111–114, 2014.
[17] W.E. Liang, and C.A. Shen, "A high performance media server and QoS routing for SVC streaming based on Software-Defined Networking," in Int. Conf. on Comput., Networking and Comm., (ICNC), pp. 556-560, 2017.
[18] M.R. Celenlioglu, and H.A. Mantar, "An SDN based intra-domain routing and resource management model," in IEEE Int. Conf. on Cloud Eng., pp. 347-352, 2015.
[19] A Mirchev, "Survey of concepts for QoS improvements via SDN," in Future Internet (FI) and Innovative Internet Technologies and Mobile Communications (IITM), Summer Semester, Munich, Germany, pp. 33-39, 2015.
[20] Y. Lu, B. Fu, X. Xi, Z. Zhang, and H. Wu, "An SDN-based flow control mechanism for guaranteeing QoS and maximizing throughput," Wireless Personal Comm., vol 97, pp.417-442, 2017.
[21] D. D'souza, K.P. Sundharan, S. Lokanath, V. Mittal, L. Perigo, and R. Hagens, Improving QoS in a Software-Defined Network, University of Colorado, 2016.
[22] R. Thenmozhi, T. Preethi, S. Sadhana, V. Shruthi, and B. Yuvarani, "Integrating multimedia services over software defined networking," Int. Res. J. Eng. and Tech (IRJET), vol. 4, 2017.
[23] L. Han, S. Sun, B. Joo, X. Jin, and S. Han, "QoS-aware routing mechanism in OpenFlow-enabled wireless multimedia sensor networks," Int. J. Dist. Sensor Networks, vol 12, p.9378120, 2016.
[24] B. Bhattacharya, and D. Das, "QoS enhancement using embedded agent in OpenfLow based SDN node," in Elect., Comp. and Comm Techn. (ICECCT), IEEE Int. Conf. on, pp. 1-5, 2015.
[25] Y. Jinyao, Z. Hailong, S. Qianjun, L. Bo, and G. Xiao, "HiQoS: An SDN-based multipath QoS solution," in Comms, China, pp. 123–133, 2015.
[26] I.F. Akyildiz, S.C. Lin, and P. Wang, "Wireless software-defined networks (W-SDNs) and network function virtualization (NFV) for 5G cellular systems: An overview and qualitative evaluation," Computer Networks, vol 93, pp.66-79, 2015.
[27] S. Civanlar, M. Parlakisik, A.M. Tekalp, B. Gorkemli, B. Kaytaz, E Onem, "A qos-enabled openflow environment for scalable video streaming," in IEEE Globecom Workshops, pp. 351-356, 2010.
[28] H.E. Egilmez, B. Gorkemli, A.M. Tekalp, and S. Civanlar, "Scalable video streaming over OpenFlow networks: An optimization framework for QoS routing," in 2011 18th IEEE International Conference on Image Processing, pp. 2241-2244, 2011.
[29] H.E. Egilmez, S.T. Dane, K.T. Bagci, and A.M.. Tekalp, "OpenQoS: An OpenFlow controller design for multimedia delivery with end-to-end Quality of Service over Software-Defined Networks,". in Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference, pp. 1-8, 2012.
[30] N. Kumar, "Mobile ad hoc network: issue and challenges related to QoS and solutions," Int. J. Eng. Technol. Sci. Res, vol 4, pp.415-419, 2017.
[31] M. Masip-Bruina, J. Yannuzzib, A. Domingo-Pascuala, M. Fonteb, E. Curadob, F. Monteirob, "Research challenges in QoS routing," Comp. Comms, vol 29, pp.563-581, 2006.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 44-56 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Improvement of the Natural Self-Timed Circuit Tolerance to Short-Term Soft Errors

Yuri Afanasyevich Stepchenkov[*,1], Anton Nikolaevich Kamenskih[2], Yuri Georgievich Diachenko[1], Yuri Vladimirovich Rogdestvenski[1], Denis Yuryevich Diachenko[1]

[1]*Department of architectural and schematic basis of the innovative computational systems, Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (IPI FRC CSC RAS), Moscow, 119333, Russian Federation*

[2]*Department of Automation and remote control, Perm National Research Polytechnic University, Perm, Russian Federation*

ARTICLE INFO

Article history:
Received: 15 October, 2019
Accepted: 01 February, 2020
Online: 09 March, 2020

Keywords:
Self-timed circuit
Soft error tolerance
Indication
Dual-rail
Anti-spacer
DICE-approach

ABSTRACT

The paper discusses the features of the implementation and functioning of digital self-timed circuits. They have a naturally high tolerance to short-term single soft errors caused by various factors, such as nuclear particles, radiation, and others. Combinational self-timed circuits using dual-rail coding of signals are naturally immune to 91% of typical soft errors classified in the paper. The remaining critical soft errors are related to the state of the dual-rail signal, opposite to the spacer and forbidden in traditional dual-rail coding of signals. Paper proposes to consider this state as the second spacer and to indicate it as a spacer to increase the self-timed circuit tolerance to soft errors. Together with an improved indication of the self-timed pipeline, this provides masking of 100% of the considered typical soft errors in combinational self-timed circuits. Due to internal feedback, self-timed latches and flip-flops are less protected from soft errors, as are synchronous memory cells. But thanks to their indication and the input and output signals generation discipline, they are also immune to 89% of typical soft errors. Usage of the self-timed latches and flip-flops with dual-rail coding of information outputs increases the tolerance of self-timed latches and flip-flops to soft errors by 2%. Application of the DICE-like approach to circuitry and layout design of sequential self-timed circuits provide an increase in their tolerance to the single soft errors up to the level of 100%.

1. Introduction

The failure-tolerance of digital integrated circuits depends, firstly, on their complexity and aging mechanisms of their components, and secondly, on the impact of a number of the destabilizing factors (noise on power and signal nets, radiation, heavy charged particles (HCP), protons, neutrons traveling through the semiconductor bulk, and others). The first type of factors leads to catastrophic failures of the CMOS integrated circuits. The second type of factors can also lead to catastrophic failures as the latch-up effect. Still, much more often, they cause short-term soft errors.

This paper discusses the natural properties of self-timed (ST) circuits, providing their high tolerance to soft errors, and proposes methods for increasing this tolerance. It is an extension of work initially presented in the 10th IEEE International Conference on Dependable Systems, Services, and Technologies, DESSERT'2019 [1].

The Single Event Transient (SET) of the digital cell (which in combinational CMOS circuits is a short-term event) is the most common [2]. A change of the memory cell state (Single Event Upset, SEU) is also possible. If the last event occurs in the memory cell storage phase, it becomes permanent rather than short-term. The duration of the failure depends on the cause of it. In this article, we consider failures caused by the passage of a nuclear particle. Depending on the particle's energy, they last up to units of nanoseconds [3].

* Yuri Afanasyevich Stepchenkov, Ystepchenkov@ipiran.ru

There are various circuitry methods for protection against soft errors. For example, Dual Data Stream Logic memory cells [4], two-phase logic cells [5], Dual Interlocked Storage Cell [6], combinational parts duplication [7], doubling each transistor in a schematic circuit [8], detection and fault isolation [9], applying additional gates [10], and so on. These solutions use the doubled or redundant implementation of circuit logical functions. There are many known techniques for improving the fault tolerance of synchronous and asynchronous circuits, for example, [11, 12], and even at the circuit synthesis level, for example, [13].

ST circuits [14 – 18] are alternative to synchronous circuits. They are initially hardware redundant, as the failure-resistant synchronous circuits are. The first reason for such redundancy is the usage of redundant code, which is mainly dual-rail one with a spacer. Secondly, ST circuits indicate the outputs of all their cells and implement a two-phase operation discipline. Such a circuit provides an increased soft-error tolerance in comparison with synchronous and even asynchronous counterparts.

The purpose of this paper is to evaluate the natural ST circuit tolerance to soft errors and suggest ways to increase it. The ST circuits analyzed in the paper belong to the quasi delay insensitive (QDI) class. They use dual-rail coding of the information signals and two-phase operation. They indicate the switching completion of all fired cells of the ST circuit in both phases of its operation. The articles and studies known to us do not contain, to our knowledge, the results of a quantitative analysis of QDI circuit tolerance to soft errors. The results of this analysis presented here are the main contribution of the paper.

The scientific novelty of this paper consists of two ideas.

The first idea is to indicate a usually forbidden state of the dual-rail signal, which is opposite to its spacer, as the correct spacer. As a result, this forbidden state becomes non-critical and does not corrupt processed data.

Another idea is to use RS-latches and RS-flip-flops with dual-rail output with spacer for storing any data. They improve ST sequential circuit tolerance to soft errors.

2. Features of ST circuits

Dual-rail coding replaces each information signal X with a dual-rail signal (DRS) {X, XB}. Usually, DRS has two working states (“01” and “10”) and one spacer. Spacer can be either zero (“00”) or a unit (“11”). The state opposite to spacer, called “anti-spacer” (AS), is prohibited in traditional ST circuits. Considering AS never appears, the classic ST indication interprets it as a valid working state (VWS). So, when it appears, for example, as a result of SET, it propagates along the circuit and corrupts data.

In addition to DRS, the ST circuits use single-rail signals – indication signals and control ones. Indication signals are generated by indication subcircuits and acknowledge the completion of the ST circuit switching into a working or spacer phase. Indication subcircuits consist of C-elements or hysteretic triggers (H-triggers) [16]. They regulate the interaction between ST parts of a total ST circuit. Therefore, we have to consider the impact of soft errors on the operation of the ST circuits with DRS and single-rail signals, which are internal and output signals for this circuit. The inputs of this ST circuit are always the output signals of some other one.

Request-acknowledge procedure for units connected by information signals (handshake procedure) accompanies a two-phase operation discipline. Each ST circuit allows switching sources of its inputs to a working (or spacer) phase only itself completes turning to a spacer (or working) phase [16]. Therefore, the ST circuit discipline masks a short-term soft error, which duration is less than the time for forming a new working state at the ST circuit outputs. Therefore, soft error criticality for the ST circuit operation depends on the time when the failure occurs within the operation cycle of the ST circuit. However, in high-performance digital ST circuits, the time factor ceases to play a decisive role.

ST sequential circuits (STSC) consist of bistable cells (BSC), which are analogs of synchronous RS-latches with cross-connections. BSC outputs form a signal having two static working states and one dynamic transit state through which the BSC passes when switching between static working states. At the absence of soft errors, its outputs at any time are in the working state (“10” or “01”) or transit state (“00” for the BSC on NOR cells or “11” for the BSC on NAND cells). The inverse state of the transit state (let's call it “anti-transit state,” ATS) never occurs during regular operation.

A soft error occurred in an STSC can lead to irreversible consequences if it causes switching BSC to an opposite working state. In contrast to the combinational circuit, BSC has positive feedback due to cross-connections. Therefore, when BSC inputs are inactive, a soft error in one half of BSC can force switching its second half as well. As a result, the BSC latches an invalid working state (IWS) and can't return itself to a VWS.

The difference between STSCs and combinational ST circuits also lies in their indication. STSC indication checks if BSC outputs and its inputs match in a working phase. Therefore, not all soft errors are critical.

As a result of soft errors in the source circuit, the following conditions may appear on the DRS inputs of the combinational ST circuit, not corresponded to the ST signal discipline:

- AS at one or more DRS inputs in a working or spacer phase,
- A spacer at one or more DRS inputs in a working phase,
- A working state at one or more DRS inputs in a spacer phase,
- IWS at one or more DRS inputs in a working phase.

Their duration depends on the time of natural or forced elimination of both factors:

- Soft error cause (e.g., the noise on communication wires and power stripes),
- Physical effects of the failures (e.g., an ionization current of the excess carriers injected by the traveled particles or radiation in a semiconductor bulk).

Traditional ST methodology considers AS is a prohibited state: it can never appear during the regular ST circuit operation. Therefore, the traditional ST circuit indicators do not expect it during both phases of its operation. However, a soft error can lead to the appearance of an AS state that violates the fundamental DRS generation discipline. The AS can propagate along the data processing path if one does not mask or stop it.

In a working phase, a DRS source can switch to VWS or AS, when AS appears at its inputs. For example, DRS with zero spacer, described by a pair of logical functions:

$$Y = A + B,$$

$$YB = AB * BB,$$

switches either to the AS $\{Y=YB=1\}$ at $\{B=0, BB=1\}$, or to the expected working state $\{Y=1, YB=0\}$ at $B=1, BB=0$, when the AS $\{A=AB=1\}$ appears at the inputs.

Premature switching one or more DRS inputs to the spacer in the working phase does not lead to negative functional consequences. In the worst case, the internal signals and ST circuit outputs can, after switching to the working state, again return to the spacer for a time corresponding to the soft error duration. At this, data remains uncorrupted, and the indication output does not "rattle." But some cells in the circuit can switch three times during the working phase. Initially, they switch to the expected working state before the failure, and then back to the spacer at the soft error, and again to the expected working state after the failure ends. Collector (e.g., H-trigger or C-element) combined these failure cell indicators can also switch three times during the ST circuit operation phase.

Premature switching DRS back to the working state during the transition to the spacer phase does not cause any negative functional consequences if it corresponds to the working state in the previous working phase. Otherwise, there is a possibility of switching ST circuit outputs to an IWS, after they have already switched to the current working state acknowledged by their indicator. (Remind of that in ST circuits, the previous unit does not turn to a spacer until the subsequent unit completes switching to a working phase.)

IWS in the working phase is the least probable. It involves the simultaneous switching of both DRS parts to an opposite value, i.e., the simultaneous appearance of failure in two different logical cells driving this DRS. However, such a soft error is the most dangerous for the ST circuit operation, as the circuit actively reads and processes its input data. The circuit may, in turn, generate an IWS at its outputs.

The traditional ST circuit indicator does not mask critical soft errors of BSC outputs and the accompanying control signals, which are information and indication outputs of the STSC. So, such errors can corrupt processed data and disturb the handshake between ST units.

3. Soft errors classification in ST circuits

Let's consider the ST circuit as a functionally complete ST unit having DRS inputs and outputs and an indication output. An external source generates DRS inputs. The indication output acknowledges the completion of switching ST circuit into the current phase. Let's assume that the ST circuit working phase lasts for the time indication output acknowledges the completion of switching this ST circuit into the spacer. Conversely, the ST circuit spacer phase corresponds to the working state of its indication output. Then the DRS can be both in a working state and in a spacer state for some time in each ST circuit phase.

In this regard, we propose a classification of all typical soft errors in the ST circuits, shown in Figure 1. All encountered observed effects can occur in both a working phase and a spacer phase of the ST circuit. A working state into which the DRS turns because of soft error may be either the expected VWS or IWS.

An ionization pulse current generated in a semiconductor body by HCP, proton, or neutron with high enough energy is one of the typical physical soft error causes. The effective track diameter of the particle traveled through the semiconductor and induced the electron-hole pairs does not exceed one micrometer (https://habr.com/post/189066 (in Russian)). In 65-nm and below CMOS process, this corresponds to the area covering the drains and sources of the homogeneous or complementary transistors in a few logical cells. Figure 2 shows a 65-nm layout fragment, including four NOR2 cells. The plan is symmetrical. CMOS n-transistors occupy the center; p-type transistors are located at left and right. Dotted circles A1 - A4 show the effective track diameter possible positions.

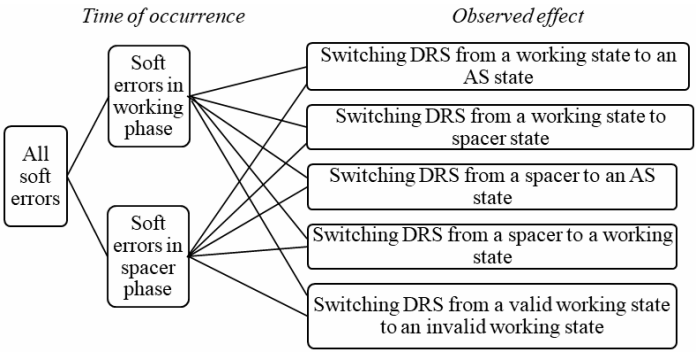


Figure 1 – Classification of the soft errors in combinational ST circuits

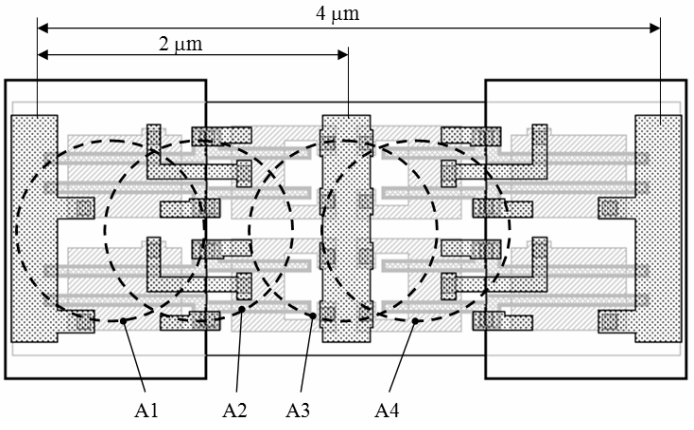


Figure 2 – Effective track diameters of the particles in the 65-nm layout

One particle cannot selectively hit, for example, only n-type transistors in one cell and only p-type transistors in another cell. It always impacts either the homogeneous transistor drains (A1, A3, A4), or the complementary transistor drains (A2) in one or

several adjacent cells at the same time. Consequently, the particles affect the structure of adjacent cells symmetrically.

Thus, one traveling particle can cause ionization currents in several adjacent cells. These currents have the same direction in all impacted cells. Therefore, the potential changes at the outputs of adjacent cells have the same polarity. In the layouts with lower design rules, the effective track diameter of the impacting particle further covers the transistor drains of both types in several adjacent cells.

If the logical cells driving DRS are forcibly placed close enough to each other on a layout, the impact of particle is symmetrical. It does not cause switching these cell outputs in opposite directions. Consequently, in the CMOS ST circuits manufactured by the 65-nm process and below, existing VWS practically never turns to IWS because of any soft error.

The ST circuit response to soft errors depends on the circuit type: combinational or sequential.

The STSC consists of one (latch) or two (flip-flop) consecutive BSC, an indication subcircuit, and BSC outputs convertors into a DRS (optional). The variety of the STSCs [19] makes difficulties for analyzing all their types in one paper. So, let's consider an STSC tolerance to soft errors by the example of ST RS-flip-flop (RSFF) shown in Figure 3. It has a DRS input {R, S} with a unit spacer, reset input (Rst), BSC output {Q, QB} with a zero-transition state (Q=QB=0), indication output (I), and consists of BSC-1, BSC-2, and indicator. BSC-1 has a unit transit state (U=UB=1).

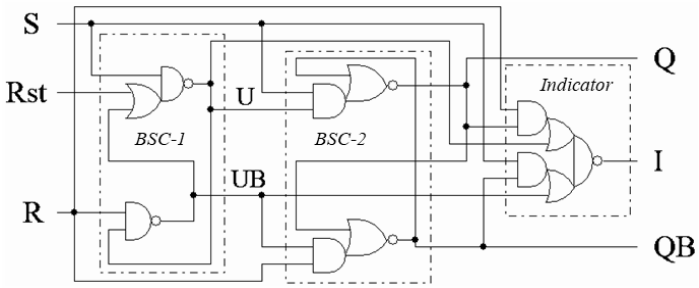


Figure 3 – ST RS-flip-flop

Figure 4 shows a classification of possible typical soft errors in ST RSFF. It includes soft errors that occur in both BSC and the indicator. For the reasons discussed above, soft error cannot lead to simultaneous switching both parts of BSC in opposite directions. Therefore, the possible occurrence of IWS is not a soft error directly. Still, it can happen as a result of the BSC transit state and the spacer at the input {R, S}.

The ST RSFF current phase is a working one if the DRS {R, S} is in the spacer. In the working phase, the RSFF information output {Q, QB} updates its state. The input {R, S} working phase corresponds to the ST RSFF spacer phase. In the spacer phase, the BSC-1 updates its state, and the BSC-2 stores its state. Note that ATS in BSC-1 causes the BSC-2 to store its state.

Let's examine the possible soft errors in combinational ST circuits and RSFFs in more detail and evaluate the probability of data corruption in the ST circuit due to them. In the first approximation, we consider the possible cases in each analyzed situation to be equally probable. Let's call the reason, caused the soft error, by the term “bug.”

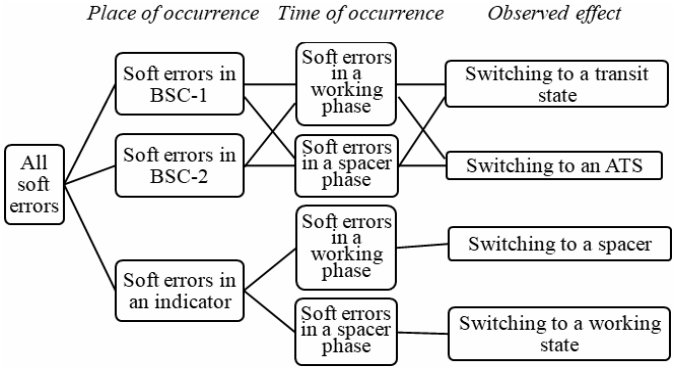


Figure 4 – Classification of the soft errors in the ST RS-flip-flops

4. Soft error cases in combinational ST circuits

According to the classification of soft errors given in Section 3 in Figure 1, the following types of soft errors are possible in a combinational ST circuit:

4.1. Switching DRS from VWS to an AS in a working phase;

4.2. Switching DRS from VWS to a spacer in a working phase;

4.3. Switching DRS from a spacer to an AS in a working phase;

4.4. Switching DRS from a spacer to a working state in a working phase;

4.5. Switching DRS from VWS to an AS in a spacer phase;

4.6. Switching DRS from VWS to a spacer in a spacer phase;

4.7. Switching DRS from a spacer to an AS in a spacer phase;

4.8. Switching the DRS from a spacer to a working state in a spacer phase.

Let's assume that any of the situations 4.1 - 4.8 may appear because of a single bug with a probability of 1/8. Let's estimate the likelihood of these soft errors corrupt data processed in the ST circuit. Corrupting data is precisely the critical situation. In other cases, a bug only leads to an additional delay in the ST circuit operation.

Case 4.1. ST circuits consider AS is a working state. Therefore, AS does not prevent the indication output of this ST circuit from switching to the working value and thereby initiating turning this ST circuit to the spacer phase. But AS actually contains a corrupted information bit. This ST circuit functioning logic does not mask AS with a probability of $P_{4.1.1} = 0.5$. AS appears at its output and then passes to the following ST circuits inputs. In the first approximation, we assume that, with a probability of $P_{4.1.2} = 0.5$, AS at the inputs of subsequent circuits corrupts their information outputs and leads to a critical error. Then the likelihood that situation 4.1 corrupts data equals to:

$$P_{4.1} = P_{4.1.1} \cdot P_{4.1.2} = 0.25.$$

Case 4.2. Switching DRS from VWS to a spacer in a working phase means that the DRS has already turned to the VWS, but a

bug caused it to switch back to the spacer. The ST circuit indication output can already acknowledge switching DRS into VWS. Since the indication subcircuit consists of triggers (i.e., C-elements or H-triggers), the indication output does not return to the spacer. With some probability, the DRS spacer causes switching of one or more DRS outputs of this ST circuit into the spacer. Subsequent ST circuits cannot complete their switch to a working phase due to the spacer on their inputs. They expect the appearance of a working state on all their significant inputs and prohibit the switching of this ST circuit into the spacer. This situation is not critical.

Case 4.3. This case means that a bug forces turning DRS to the AS instead of VWS. Similar to situation 4.1, the probability that situation 4.3 becomes critical equals $P_{4.3} = 0.25$.

Case 4.4. Switching DRS from a spacer to a working state in a working phase means that a bug caused a premature DRS transition to the working state. If this working state matches the expected VWS, the situation is not critical. However, with a probability of $P_{4.4.1} = 0.5$, this working state is IWS. As we noted earlier, the bug cannot affect both parts of DRS in opposite directions. Therefore, in this situation, a soft error appears on only one DRS part. Then the second DRS part switches to a working value under the influence of the working state of ST circuit inputs. As a result, a fail DRS turns to AS. Similar to situation 4.1, the probability that situation 4.4 becomes critical equals

$$P_{4.4} = P_{4.4.1} \cdot P_{4.1} = 0.125.$$

Case 4.5. The ST circuit has already completed the switch to a working phase. With a probability of $P_{4.5.1} = 0.5$, the functional logic of this ST circuit does not mask AS. AS appears at the output of this ST circuit, and then passes to the following ST circuit inputs. With a probability of $P_{4.5.2} = 0.5$, the ST circuit output receivers have not yet switched to a working state, and AS takes part in the generation of their outputs. Then the probability that situation 4.5 becomes critical equals to:

$$P_{4.5} = P_{4.5.1} \cdot P_{4.5.2} \cdot P_{4.1} = 0.0625.$$

Case 4.6. The ST circuit already completely switched to a working phase, and this soft error means a premature spacer. With some probability, the functional logic of this ST circuit does not mask the premature spacer. It appears at the output of this ST circuit. If these output receivers have not yet switched to a working state, a premature spacer can lock this switching. ST receivers suspend their work in anticipation of a working state on the failed inputs. They do not allow the inputs of this ST circuit in which a soft error occurred to switch to a spacer. This situation is not critical.

Case 4.7. The ST circuit has completed switching to a working phase, and DRS in which the soft error occurred has already turned to a spacer. The ST circuit output receivers also completed the switch to a working phase. AS can pause the switching this ST circuit to a spacer for the bug duration but is not critical.

Case 4.8. The ST circuit and its output receivers completed switching to the working phase, and DRS, in which the soft error occurred, has already turned to a spacer. The working state is premature. It can pause the switching this ST circuit to a spacer for the bug duration but is not critical.

Thus, only the soft errors in cases 4.1, 4.3, 4.4, and 4.5 are critical for the combinational ST circuit operation. The total probability of corrupting data processed in the ST circuits, because of soft errors classified in Section 3, equals to:

$$P_4 = (1/8) \cdot (P_{4.1} + P_{4.3} + P_{4.4} + P_{4.5}) = 0.0859375 \approx 8.6\%.$$

The soft errors that appeared in the indication subcircuit of the ST circuits are not critical. They can only cause an additional delay in the switching circuit from one phase to another.

5. Soft error cases in RSFFs

According to the classification of soft errors in the RSFFs in Figure 4, the following their types are possible:

5.1. Switching BSC-1 to a transit state in a working phase;

5.2. Switching BSC-1 to an ATS in a working phase;

5.3. Switching BSC-1 to a transit state in a spacer phase;

5.4. Switching BSC-1 to an ATS in a spacer phase;

5.5. Switching BSC-2 to a transit state in a working phase;

5.6. Switching BSC-2 to an ATS in a working phase;

5.7. Switching BSC-2 to a transit state in a spacer phase;

5.8. Switching BSC-2 to an ATS in a spacer phase;

5.9. Switching indicator to a spacer in a working phase;

5.10. Switching indicator to a working state in a spacer phase.

Let's assume that any of the cases 5.1 - 5.10 may appear because of a single bug with a probability of 1/10. Let's estimate the likelihood that these soft errors become critical and corrupt processed data.

Case 5.1. In a working phase of the ST RS-flip-flop, Rst = 0, R = S = 1. Suppose that before the soft error, the state of BSC-1 was (U = 0, UB = 1). The BSC-1 can switch to a transit state in the RSFF working phase in two situations:

5.1.1 Bug affects a logical level of a single BSC-1 output.

5.1.2. The bug affects the logical levels of both BSC-1 outputs.

In situation 5.1.1, the transit state of BSC-1 (U = UB = 1) turns to IWS (U = 1, UB = 0) as the bug affects only the output U of BSC-1. As a result, BSC-1 stores IWS. The spacer R = S = 1 causes an immediate rewriting of this IWS to BSC-2 and corrupts an information bit stored in the RSFF and presented at its output. Data corruption, not only in the RSFF but also in the subsequent data processing path, occurs in the following cases:

- With a probability of $P_{5.1.1.1} = 0.5$, the input {R, S} remains in the spacer until switching BSC-1 into IWS and writing IWS to BSC-2. With a likelihood of $P_{5.1.1.2} = 0.5$, IWS at the output {Q, QB} corrupts the data in the RSFF output receivers,

- With a probability of $P_{5.1.1.3} = 0.5$, the input {R, S} has time to switch to a new VWS before writing IWS from BSC-1 to BSC-2. Still, the new VWS with a probability of $P_{5.1.1.4} = 0.5$ is opposite

to the IWS stored in BSC-1 and does not prevent writing IWS to BSC-2. In this case, BSC-1 returns to the transit state and supports the indicator working value. With a probability of $P_{5.1.1.5} = 0.5$, IWS at the output {Q, QB} corrupts the data in the RSFF output receivers.

In situation 5.1.2, the bug does not allow BSC-1 to switch from the transit state to the IWS. At R = S = 1, BSC-2 turns to its transit state (Q = QB = 0), which traditional ST circuits consider as a spacer, and the RSFF remains in the working phase. There are two ways of situation evolution:

- With a probability of $P_{5.1.2.1} = 0.5$, the RSFF output receivers complete switching to a working phase, despite the transit state of the RSFF output. Together with the RSFF indicator, they initiate the DRS {R, S} transition to a new VWS. With a probability of $P_{5.1.2.2} = 0.5$, BSC-1 is still in transit state. Then the VWS at RSFF input is written to BSC-2 as IWS. Still, the RSFF indicator remains in a working value and suspends the ST circuit operation until the bug ends. When bug ends, BSC-1 turns to VWS and toggles the RSFF indicator to a spacer. IWS in BSC-2 does not prevent the ST circuit from completing a transition to a spacer and initiating the switching of the RSFF input {R, S} to a spacer (R = S = 1), which rewrites VWS from BSC-1 to BSC-2. ST circuit continues operation with correct data;

- With a probability of $P_{5.1.2.3} = 0.5$, the RSFF output receivers do not complete switching to a working phase because of the RSFF output transit state. The ST circuit pauses while waiting for the RSFF working state. The DRS {R, S} remains in a spacer. At the bug end, the BSC-1 state becomes metastable. With a probability of $P_{5.1.2.4} = 0.5$, BSC-1 switches to IWS. The spacer R = S = 1 causes an immediate rewriting IWS to BSC-2. ST circuit continues to work with corrupted data.

The total data corruption probability in case 5.1 equals to:

$$P_{5.1} = 0.5 \cdot (P_{5.1.1.1} \cdot P_{5.1.1.2} + P_{5.1.1.3} \cdot P_{5.1.1.4} \cdot P_{5.1.1.5}) + \\ + 0.5 \cdot (P_{5.1.2.3} \cdot P_{5.1.2.4}) = 0.3125.$$

Case 5.2. BSC-1 can switch to ATS (U = UB = 0) in an RSFF working phase also in two situations:

5.2.1 The bug affects a logical level of a single BSC-1 output.

5.2.2. The bug affects the logical levels of both BSC-1 outputs.

Suppose that before the bug, the state of BSC-1 was (U = 0, UB = 1). ATS in BSC-1 appears due to a failed switching UB from the logical 1 level to logical 0 level. In a working phase of ST RS-flip-flop, Rst = 0, R = S = 1, and BSC-2 stores VWS. ATS in BSC-1 does not damage the BSC-2 state, but ATS is not stable.

In situation 5.2.1, at the spacer R = S = 1, BSC-1 turns to IWS (U = 1, UB = 0), that is written to BSC-2, corrupting the data in the RSFF output receivers with a probability of $P_{5.2.1.1} = 0.5$.

In situation 5.2.2, at the spacer R = S = 1, BSC-1 remains in ATS until the bug end. With a probability of $P_{5.2.2.1} = 0.5$, the RSFF output receivers do not have time to switch into a working phase. So, they do not allow DRS {R, S} source to turn the RSFF inputs into a new VWS before the bug ends. At the bug end, BSC-1 is in a metastable state. With a probability of $P_{5.2.2.2} = 0.5$, it turns to a working state opposite to BSC-2 state, which is written to BSC-2, corrupting the data in the RSFF output receivers with a probability of $P_{5.2.2.3} = 0.5$.

The data corruption probability in case 5.2 equals to:

$$P_{5.2} = 0.5 \cdot P_{5.2.1.1} + 0.5 \cdot (P_{5.2.2.1} \cdot P_{5.2.2.2} \cdot P_{5.2.2.3}) = 0.3125.$$

Case 5.3. BSC-1 can switch to a transit state (U = UB = 1) in an RSFF spacer phase in two situations:

5.3.1. The bug affects a logical level of a single BSC-1 output.

5.3.2. The bug affects the logical levels of both BSC-1 outputs.

In situation 5.3.1, BSC-1 may switch from the transit state to IWS. Let, for example, R = Rst = 0, S = 1, U = 0, UB = 1. BSC-1 output U switches to U = 1 because of the bug. With a probability of $P_{5.3.1.1} = 0.5$, the RSFF and its output receivers have already completed the transition to a spacer and initiated switching DRS {R, S} to a spacer. BSC-1 remains in transit state until input R switches to a spacer value (R = 1). With a probability of $P_{5.3.1.2} = 0.5$, the bug does not end by this time, and BSC-1 switches from the transit state to IWS (U = 1, UB = 0). IWS is written from BSC-1 to BSC-2 since R = S = 1 and appears at the RSFF output. With a probability of $P_{5.3.1.3} = 0.5$, IWS corrupts the data in the RSFF output receivers.

In situation 5.3.2, BSC-1 may also switch from transit state to IWS. Let, for example, R = Rst = 0, S = 1, U = 0, UB = 1. The bug switches BSC-1 output U to U = 1 and reinforces UB=1 value. With a probability of $P_{5.3.2.1} = 0.5$, the RSFF and its output receivers have already completed the transition to a spacer and initiated switching DRS {R, S} to a spacer. The BSC-1 remains in transit state until input R switches to the spacer value (R = 1). With a probability of $P_{5.3.2.2} = 0.5$, the bug does not end by this time, but BSC-1 cannot switch from the transit state to any working state because the bug forces the logical 1 level at both BSC-1 outputs. As a result, upon bug completion, BSC-1 is in a metastable state. With a probability of $P_{5.3.2.3} = 0.5$, it switches to the working state opposite to that of BSC-2. BSC-2 turns into this state. With a probability of $P_{5.3.2.4} = 0.5$, it corrupts data in the RSFF output receivers.

The total data corruption probability in case 5.3 equals to:

$$P_{5.3} = 0.5 \cdot (P_{5.3.1.1} \cdot P_{5.3.1.2} \cdot P_{5.3.1.3}) + \\ + 0.5 \cdot (P_{5.3.2.1} \cdot P_{5.3.2.2} \cdot P_{5.3.2.3} \cdot P_{5.3.2.4}) = 0.09375.$$

Case 5.4. BSC-1 can switch to ATS (U = UB = 0) in the RSFF spacer phase also in two situations:

5.4.1. The bug affects a logical level of a single BSC-1 output.

5.4.2. The bug affects the logical levels of both BSC-1 outputs.

In situation 5.4.1, BSC-1 switches from ATS to IWS. Let, for example, R = Rst = 0, S = 1, U = 0, UB = 1 before the bug. The output UB of BSC-1 switches to UB = 0 because of the bug. Since the bug did not affect the U node, condition UB = Rst = 0 causes switching U from logical 0 to logical 1 level, and BSC-1 turns to IWS (U = 1, UB = 0). The following writing IWS to BSC-2 turns indicator to a working value I = 0. Two directions of further evolution are possible:

- With a probability of $P_{5.4.1.1} = 0.5$, the RSFF and its output receivers have already completed the transition to a spacer and initiated switching DRS {R, S} to a spacer. With a probability of $P_{5.4.1.2} = 0.5$, BSC-1 remains in the IWS until the input R switches to a spacer value (R = 1). The RSFF input spacer writes the IWS from BSC-1 to BSC-2, which with a probability of $P_{5.4.1.3} = 0.5$ corrupts data in the RSFF output receivers;

- With a probability of $P_{5.4.1.3} = 0.5$, the RSFF and its output receivers do not have time to switch to a spacer. Input {R, S} remains in VWS until the bug end. At the bug end, BSC-1 switches to VWS, the RSFF indicator acknowledges the RSFF spacer state, and the ST circuit continues operation with the correct data.

In the situation, 5.4.2 BSC-1 can also switch from ATS to IWS. Let, for example, R = Rst = 0, S = 1, U = 0, UB = 1 before the bug. Bug switches the output UB of BSC-1 to UB = 0 and reinforce the value U = 0. BSC-1 remains in ATS and provides storage of VWS in BSC-2. The indicator acknowledges the RSFF spacer state. With a probability of $P_{5.4.2.1} = 0.5$, the RSFF output receivers complete the transition to a spacer and initiate switching DRS {R, S} to a spacer. With a probability of $P_{5.4.2.2} = 0.5$, BSC-1 remains in ATS until the input R switches to a spacer value (R = 1). Until the bug end, BSC-1 cannot switch from ATS to any working state, because the bug holds logical 0 levels at both outputs of BSC-1. At the bug end, BSC-1 enters a metastable state, which with a probability of $P_{5.4.2.3} = 0.5$, switches into the working state opposite to the initial VWS of DRS input {R, S}. BSC-2 turns to this IWS, and with a probability of $P_{5.4.2.4} = 0.5$, corrupts data in the RSFF output receivers.

The total data corruption probability in case 5.4 equals to:

$$P_{5.4} = 0.5 \cdot (P_{5.4.1.1} \cdot P_{5.4.1.2} \cdot P_{5.4.1.3}) + \\ + 0.5 \cdot (P_{5.4.2.1} \cdot P_{5.4.2.2} \cdot P_{5.4.2.3} \cdot P_{5.4.2.4}) = 0.09375.$$

Case 5.5. In a working phase of an RSFF, Rst = 0, R = S = 1. Suppose that before a bug, the states of the BSCs were (U = 0, UB = 1), and (Q = 1, QB = 0). BSC-2 can switch to a transit state in the RSFF working phase in two situations:

5.5.1. The bug affects a logical level of a single BSC-2 output.

5.5.2. The bug affects the logical levels of both BSC-2 outputs.

In situation 5.5.1, the transit state of BSC-2 (Q = QB = 0) switches the RSFF indicator to a spacer. Further situation evolution is possible in two directions:

- With a probability of $P_{5.5.1.1} = 0.5$, the RSFF output receivers complete switching into a working phase and initiate the transition of DRS {R, S} to a new VWS. A spacer value of the RSFF indicator remains until the bug ends. If a new VWS appears before the bug end, it supports the spacer value of the RSFF indicator, and the ST circuit continues regular operation. If the bug ends before the new input VWS, a condition R = S = 1 causes writing VWS from BSC-1 to BSC-2. It does not corrupt data in the RSFF output receivers;

- With a probability of $P_{5.5.1.2} = 0.5$, the RSFF output receivers do not complete switching into a working phase and do not initiate the transition of DRS {R, S} to a new VWS. The receivers consider the RSFF output transit state as a spacer. ST circuit pauses with its operation. At the bug end, the spacer R = S = 1 forces writing VWS from BSC-1 to BSC-2. BSC-2 restores its VWS, and the ST circuit continues regular operation with correct data.

Situation 5.5.2 is similar to the situation 5.5.1 and also cannot cause any data corruption. So, the total data corruption probability in case 5.5 equals $P_{5.5} = 0$.

Case 5.6. Let Rst = 0, R = S = 1, U = 0, UB = 1, Q = 1, QB = 0 in the ST RS-flip-flop working phase. BSC-2 can switch to ATS in two situations:

5.6.1. The bug affects only a logical level of QB output.

5.6.2. The bug affects the logical levels of both BSC-2 outputs.

In situation 5.6.1, ATS (Q = QB = 1) in BSC-2 forces switching BSC-2 to IWS (Q = 0, QB = 1) due to cross-connections. The RSFF indicator turns to a spacer value I = 1 and prevents the RSFF output receivers from using the {Q, QB} state.

With a probability of $P_{5.6.1.1} = 0.5$, the RSFF output receivers do not complete the switch to a working phase. They do not allow the RSFF input {R, S} source to initiate the transition of DRS {R, S} to a new VWS. At the bug end, the spacer R = S = 1 forces writing VWS from BSC-1 to BSC-2. BSC-2 restores its VWS, the RSFF indicator switches to a working value, and the ST circuit continues regular operation with correct data.

With a probability of $P_{5.6.1.2} = 0.5$, the RSFF output receivers have time to complete the switching to a working phase and allow the RSFF input {R, S} source to initiate the transition of DRS {R, S} to a new VWS. With a probability of $P_{5.6.1.3} = 0.5$, IWS at the output {Q, QB} during the switching RSFF indicator from a working value to a spacer has time to corrupt data in subsequent ST circuits. With a probability of $P_{5.6.1.4} = 0.5$, DRS {R, S} has time to switch to the new VWS before the bug ends. The RSFF indicator remains in the spacer and not allow for restoring corrupted data.

In situation 5.6.2, the ATS in BSC-2 (Q = QB = 1) does not switch into any working state, because of the bug forces logical 1 level on both BSC-2 outputs. The RSFF indicator remains in a working value I = 0 and allows the RSFF output receivers to use the {Q, QB} state. With a probability of $P_{5.6.2.1} = 0.5$, ATS at the output {Q, QB} corrupts data in the receivers of this output, since they consider it is a working state. With a probability of $P_{5.6.2.2} = 0.5$, the RSFF output receivers complete switching to a working phase. They initiate the transition of DRS {R, S} to a new VWS, which occurs, with a probability of $P_{5.6.2.3} = 0.5$, before the bug ends. The RSFF indicator switches to a spacer I = 1 and prohibits restoring corrupted data.

The total data corruption probability in case 5.6 equals to:

$$P_{5.6} = 0.5 \cdot (P_{5.6.1.2} \cdot P_{5.6.1.3} \cdot P_{5.6.1.4}) + \\ + 0.5 \cdot (P_{5.6.2.1} \cdot P_{5.6.2.2} \cdot P_{5.6.2.3}) = 0.125.$$

Case 5.7. Let's suppose that in the ST RS-flip-flop spacer phase, Rst = 0, R = 1, S = 0, U = 1, UB = 0, and before the bug, BSC-2 state was (Q = 0, QB = 1). In a spacer phase, the BSC-1 static state and the RSFF inputs do not affect the BSC-2 state.

BSC-2 can switch to the transit state (Q = QB = 0) in the RSFF spacer phase in two situations:

5.7.1. The bug affects only a logical level of QB output.

5.7.2. The bug affects the logical levels of both BSC-2 outputs.

In situation 5.7.1, the transit state of BSC-2 (Q = QB = 0) turns to IWS (Q = 1, QB = 0) due to cross-connections, since the bug does not affect the logical level of Q. But the indicator remains in a spacer, and following the handshake discipline, the RSFF output receivers do not use IWS at the output of this RSFF. The RSFF indicator and its output receivers initiate the transition of DRS {R, S} to a spacer.

With a probability of $P_{5.7.1.1} = 0.5$, VWS in BSC-1 does not correspond to the BSC-2 state, as in the case under consideration. Since bug only affects QB, the condition (S = U = 1) causes switching output Q to Q = 0. BSC-2 turns back to the transit state considered as a spacer by subsequent ST circuits. The indicator remains in a spacer value I = 1. The RSFF output receivers do not complete switching to a working phase and do not allow the RSFF input {R, S} source to initiate the transition of DRS {R, S} to a new VWS. At the bug end, the QB output switches to QB = 1, and BSC-2 turns to the expected VWS. The ST circuit continues regular operation with correct data.

If the states of BSC-1 and BSC-2 coincide, then after the DRS {R, S} transition to a spacer, the ST circuit continues regular operation with uncorrupted data.

In situation 5.7.2, the transit state of BSC-2 (Q = QB = 0) does not switch into a working state, since the bug forces the logical level of both RSFF outputs. But the indicator also remains in a spacer. Following the handshake discipline, the RSFF output receivers switch to a spacer. Together with the RSFF indicator, they initiate the transition of DRS {R, S} to the spacer. Upon bug completion and switching DRS {R, S} to the spacer, BSC-2 turns to the state of BSC-1, and the ST circuit continues regular operation with correct data.

The total data corruption probability in case 5.7 equals to $P_{5.7} = 0$.

Case 5.8. Let's suppose that in the ST RS-flip-flop spacer phase, Rst = 0, R = 1, S = 0, U = 1, UB = 0, and BSC-2 state is (Q = 0, QB = 1) before the bug. BSC-2 can switch to ATS in two situations:

5.8.1. The bug affects only a logical level of Q output.

5.8.2. The bug affects the logical levels of both BSC-2 outputs.

In situation 5.8.1, ATS in BSC-2 (Q = QB = 1) causes switching BSC-2 into IWS (Q = 1, QB = 0) due to cross-connections. The indicator is in a spacer value I = 1 regardless of the BSC-2 state. It prevents the RSFF output receivers from using the state {Q, QB}. The RSFF output receivers complete switching to a spacer and allow the RSFF input {R, S} source to initiate the transition of DRS {R, S} to a spacer.

With a probability of $P_{5.8.1.1} = 0.5$, DRS {R, S} has time to switch to the spacer (R = S = 1) before the bug ends. With a probability of $P_{5.8.1.2} = 0.5$, VWS in BSC-1 does not correspond to the BSC-2 state, as in the case under consideration. Since the bug forces logical 1 level at the output Q, condition (S = U = 1) cannot switch the output Q to Q = 0. BSC-2 remains in IWS, but the RSFF indicator does not turn to a working value I = 0. It does not allow the RSFF output receivers to use the {Q, QB} state before the bug ends, and BSC-2 turns into VSC stored in BSC-1. Only then the RSFF indicator switches to a working value I = 0 and allow the RSFF output receivers to use the state {Q, QB}. The ST circuit continues regular operation with correct data. If, after switching DRS {R, S} to the spacer, the VWS in BSC-1 coincides BSC-2 state, the ST circuit also continues to work with the correct data.

With a probability of $P_{5.8.1.3} = 0.5$, DRS {R, S} does not have time to switch to the spacer (R = S = 1) before the bug ends. The input {R, S} working state, and VWS in BSC-1 prevent writing to the BSC-2. At the bug end, BSC-2 continues to store IWS. Still, due to the spacer value of the RSFF indicator, the RSFF output receivers do not use it, and their data remains uncorrupted.

In situation 5.8.2, ATS in BSC-2 (Q = QB = 1) does not cause the BSC-2 to switch into any working state, because of the bug forces logical 1 level on both outputs of BSC-2. The RSFF indicator remains in a spacer value I = 1 and prevents the RSFF output receivers from using ATS. The RSFF output receivers complete switching to the spacer and allow the RSFF input {R, S} source to initiate the transition of DRS {R, S} to a spacer.

With a probability of $P_{5.8.2.1} = 0.5$, DRS {R, S} has time to switch to the spacer (R = S = 1) before the bug ends. Because of ongoing bug, BSC-2 does not turn into VWS stored in BSC-1. ATS at the output {Q, QB} causes switching the RSFF indicator to a working value I = 0, which allows the RSFF output receivers to use the ATS as a working state. With a probability of $P_{5.8.2.2} = 0.5$, the ATS corrupts data in the RSFF output receivers.

With a probability of $P_{5.8.2.3} = 0.5$, DRS {R, S} does not have time to switch to the spacer (R = S = 1) before the bug ends. The input {R, S} working state, and VWS in BSC-1 prevent writing to BSC-2. After the bug ends, BSC-2 is in a metastable state, which switches to an arbitrary working state. But due to the spacer value of the RSFF indicator, the output {Q, QB} working state is not used by the RSFF output receivers, and their data remains uncorrupted.

The total data corruption probability in case 5.8 equals to:

$$P_{5.8} = 0.5 \cdot (P_{5.8.2.1} \cdot P_{5.8.2.2}) = 0.125.$$

Case 5.9. It suggests the indicator has switched to the working value but returned to the spacer because of the bug. In this case, BSC-1 and BSC-2 store VWS. Since the data inside and at the RSFF information output do not change, a failed switching indicator to the spacer, in the worst case, only pauses the ST circuit regular operation. When the bug ends, the RSFF indicator returns to the working value, and the ST circuit continues the correct operation.

The total data corruption probability in case 5.9 equals to $P_{5.9} = 0$.

Case 5.10. It assumes the indicator has switched to the spacer but returned to the working state because of the bug. In this case, BSC-1 and BSC-2 store VWS. Similarly, to case 5.9, when the

bug ends, the RSFF indicator returns to the spacer, and the ST circuit continues the correct operation.

The total data corruption probability in case 5.10 equals to $P_{5.10} = 0$.

The reasoning given above for the specific values of the inputs, outputs, and internal state of the ST RS-flip-flop in cases 5.1 - 5.8 is also valid for any other initial conditions.

Thus, only the soft errors in the ST RS-flip-flop in cases 5.1 - 5.4, 5.6, and 5.8 are critical for the ST circuit functioning. The total data corruption probability because of a single soft error in the ST RS-flip-flop equals to:

$$P_5 = (1/10) \cdot (P_{5.1} + P_{5.2} + P_{5.3} + P_{5.4} + P_{5.6} + P_{5.8}) = 0.10625.$$

Failure analysis for other types of STSCs gives a similar result. We do not cite it due to the limited volume of the paper.

6. Masking critical failures in ST circuits

6.1 Masking soft errors in ST combinational circuits

The following two methods, being used simultaneously, can improve soft error tolerance of the ST circuit:

- masking AS state,
- improvements in the ST pipeline indication.

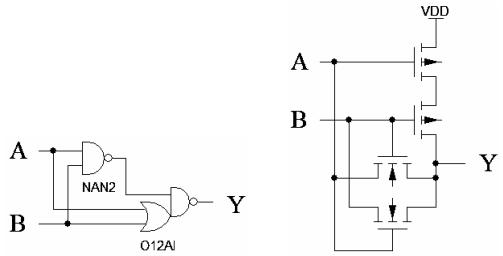
The first method uses the fail-safe DRS discipline, considering AS is the second spacer [1], as shown in Table 1. Two spacers inside the same circuit were used earlier in secure dual-rail logic [20] and NCL circuits [11]. There they were formed by the different DRS sources and provided increased security for data encoding [20] or reduced power consumption [11]. We propose to indicate the AS as a valid second spacer of the same DRS.

Table 1: Fail-safe DRS discipline in ST circuits

S.No	Synchronous signal X	ST signal		State
		X	XB	
1	0	0	1	bit 0
2	1	1	0	bit 1
3	-	0	0	spacer 0
4	-	1	1	spacer 1

XNOR, or “equivalence” cell, indicates two DRS spacers (“00” and “11”). Figure 5 shows two XNOR implementations suitable for use in QDI circuits. The circuit in Figure 5(a) [21] is convenient for implementation with a standard cell library. The circuit in Fig. 1(b) has the least possible transistor number.

The two-phase discipline of the ST circuit operation and the requirement of a mandatory indication of each DRS mask a mixed spacer-working state at the circuit inputs. Indeed, the DRS consistently switches from spacer to a working state and back to spacer state. The indicator acknowledges the completion of this switching to each state. The indicator also acknowledges the incorrect switching DRS to a state that does not correspond to the ST circuit operation phase. As a result, the indicator does not correspond to an expected one in the current phase of the ST circuit. An indication subcircuit detects and localizes this case.



(a) (b)

Figure 5 – XNOR implementation: on standard cells (a) and pass CMOS transistors (b)

Thus, the usage of a double spacer allows one to mask all soft errors causing the AS appearance in combinational ST circuits, namely, cases 4.1, 4.3, 4.4, and 4.5 in Section 4.

As a result, with the indication subcircuit slight complication providing AS indication as a spacer, the combinational ST circuits mask soft errors in all cases analyzed in Section 4. An improved method for controlling the phases of the output register during pipeline implementation of the ST circuit retains this advantage within the framework of the general ST circuit.

Figure 6 shows a circuit of an optimized ST pipeline utilized in the Fused Multiply-Add (FMA) project [22]. The logic and register parts implement the pipeline stage. LI* and RI* blocks compress the logical part and register bitwise indicators in each pipeline stage into one total indication signal, respectively. The output signals of the LI* and RI* blocks indicate the completion of the switching of all cells in the pipeline stage, “fired” by the stage inputs in the current operation phase.

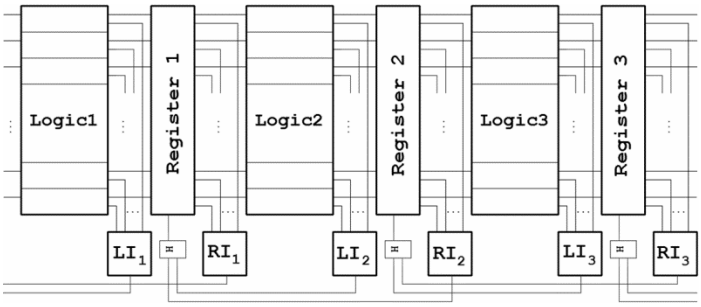


Figure 6 – Traditional ST pipeline

H-trigger [16] collects the outputs of the blocks LI_{k+1} and RI_{k+1}, forming a control signal for the output register of the k-th pipeline stage within the ST pipeline handshake discipline. The k-th stage output register, in turn, forms the working state and the spacer of the logical part inputs of the (k+1)-th pipeline stage.

The H-trigger is a static analog of the Muller C-element [16], traditionally used for ST circuit indication. Figure 7 shows the two-input H-trigger CMOS circuit. Unlike the C-element, the H-trigger does not contain a “weak” inverter, and due to this, it has better noise immunity. Table 2 represents the truth table of the two-input H-trigger. If the states of the inputs I0 and I1 are equal, then output Q turns to the same state. Otherwise, output Q stores its state.

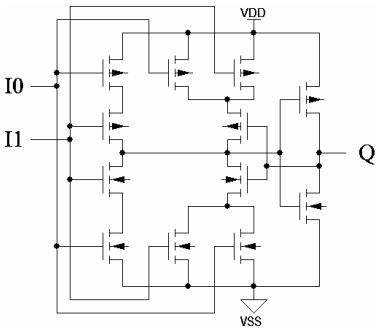


Figure 7 – CMOS circuit of the two-input H-trigger

Table 2: Truth table of the two-input H-trigger

S.No	Inputs		Output, Q^+
	I0	I1	
1	0	0	0
2	1	0	Q
3	0	1	Q
4	1	1	1

Note that the pipeline control circuit in Figure 6 is correct only if the logic block of the current pipeline stage indicates all previous pipeline stage register outputs. Otherwise, the H-trigger controlling the k-th stage register must have a third input connected to this register indication output.

Figure 8(a) shows a one register bit implementation, where the DRS inputs $\{X_j, XB_j\}$ have a unit spacer. It consists of two H-triggers. Due to this, it stores both the working and spacer states of DRS inputs $\{X_j, XB_j\}$, while ensuring an indication of all its inputs and outputs. Ph input is a total phase control signal for all register bits. When the k-th pipeline stage combinational part switches to a spacer phase, its output register, for some time, is in the working phase and provides the k-th stage outputs to the (k+1)-th stage inputs.

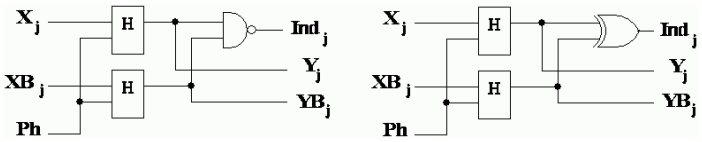


Figure 8 – ST register bit: traditional (a) and soft error resisted (b)

The indicators of the combinational part and the register of the (k+1)-th stage acknowledge their switching to the working phase. Then the H-trigger combining them forms the spacer value of the general register control signal Ph. Only after this, the register becomes insensitive to the working state changes of the inputs $\{X_j, XB_j\}$. Therefore, for some time, any working state change at the output of the k-th stage combinational part is stored in its output register and transferred to the (k+1)-th pipeline stage inputs.

The register bit circuit complication, as shown in Figure 8 (b), increases its tolerance to soft errors. Then register indicates the AS state of the $\{Y_j, YB_j\}$ output as a spacer.

Table 3 displays the data corruption probabilities for the cases discussed in Section 4 in the classical combinational ST circuit and the ST circuit using the proposed circuit techniques increasing soft error tolerance. The last row shows the total data corruption probability in combinational ST circuits because of soft errors analyzed in Section 4, taking into account the equal likelihood of their occurrence.

Table 3: Data corruption probabilities in classical and improved combinational ST circuit

Soft error case	Classical ST circuit	Improved ST circuit
4.1	0.25	0
4.2	0	0
4.3	0.25	0
4.4	0.125	0
4.5	0.125	0
4.6	0	0
4.7	0	0
4.8	0	0
Total	8.6%	0%

Thus, classical ST combinational circuits mask soft errors classified in Section 3 in 91.4% of their occurrence cases. AS indication as a spacer and the bit implementation of the pipeline stage output register by the circuit in Figure 8(b) mask the remaining 8.6% of soft error cases in the combinational ST circuit that are not masked by the classical ST circuits properties and their interaction discipline.

Remind, we consider the soft errors that occur too close to the ST circuit current phase end or having a too long duration. So, the ST circuit two-phase discipline cannot mask them. In synchronous circuits, critical soft errors having no time to disappear before the active clock edge arrives. They can only be masked by using fail-safe coding or through special circuitry techniques that increase the hardware complexity.

6.2 Masking soft errors in STSCs

The usage of ST RSFFs with DRS information outputs [19] and circuitry methods improving the failure-tolerance of memory circuits (for example, LTMR [23], DICE [24, Figure 3]) allow for essential increasing the STSC soft error tolerance.

Combinational ST circuits use DRS inputs, while the RSFF information outputs in Figure 3 form a signal without spacer. The RSFF with DRS information input {R, S} and output {QP, QPB}, shown in Figure 9, resolves this problem.

The RSFFs with DRS information output write the BSC-2 state to their information outputs only in the working phase of the BSC-2 and only when its state corresponds to the BSC-1 state. The inputs of the RSFF shown in Figure 9 have a unit spacer (R = S = 1), which forces writing BSC-1 state to BSC-2. Indicator value I = 0 acknowledges this write successful completion and allows for switching DRS outputs (QP, QPB) to a working state corresponding to the BSC-2 state. The working state at the RSFF information inputs ({R = 0, S = 1} or {R = 1, S = 0}) causes switching the RSFF indication output to I = 1 and forces turning DRS output to the spacer (QP = QPB = 0). Therefore, any change in the BSC-2 state in this phase does not affect the RSFF output state.

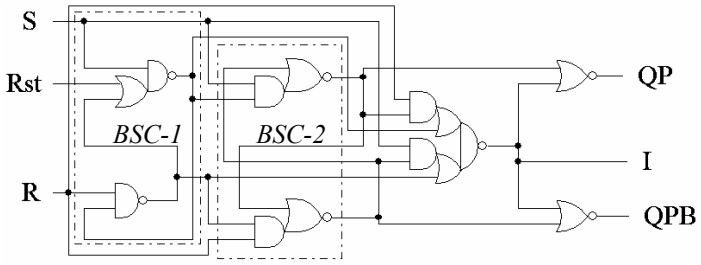


Figure 9 – ST RS-flip-flop with DRS input and output

The indication output spacer value I = 1 or ATS in BSC-2 force the spacer at the {QP, QPB} output of the RSFF in Figure 9. As a result, the data corruption probabilities $P_{5.6.2.1}$ and $P_{5.8.2.2}$ for ATS at {QP, QPB} output become zero. Then the probabilities of cases 5.6 and 5.8 from Section 5 become equal to $P_{5.6} = 0.0625$ and $P_{5.8} = 0$, and the total data corruption likelihood because of the soft error in the RSFF shown in Figure 9 decreases to $P_5 = 0.0875 \approx 8.8\%$.

DICE (Dual Interlocked Cell) is a particular circuitry method for protecting RAM cells and a sequential logic against soft errors. First, it doubles hardware, using cross-connections to control the gates of p- and n-type transistors in the symmetric parts of the circuit. Second, it assumes spacing the corresponding layout fragments to avoid a simultaneous impact on them the same particle. Figure 10 shows an example of a dice-style synchronous D-latch [24, Figure 3].

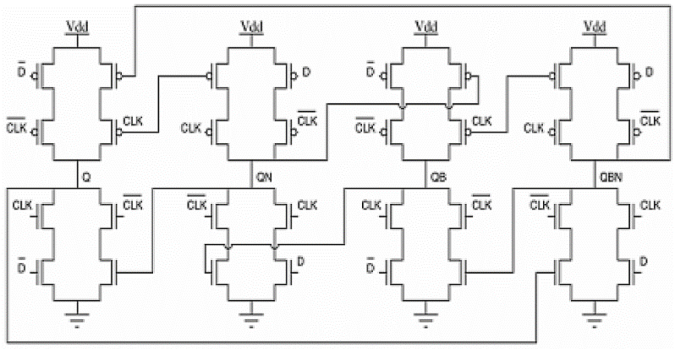


Figure 10 – Synchronous DICE-style D-latch

LTMR (Localized Triple Modular Redundancy) is a method of local majoring applied at the level of individual logical cells. Figure 11 demonstrates an example of a synchronous LTMR flip-flop [23].

The LTMR method surpasses the DICE approach in protecting high-performance synchronous circuits against single failures. It masks failure in the assumption that the failure occurs only in one of the three parts of the circuit and does not corrupt the result of data processing. Unlike LTMR, DICE prevents writing incorrect state to the sequential cell and provides a self-healing of the circuit after the end of the failure. But writing the correct state to the sequential cell requires a time exceeding the duration of the failure. Therefore, the effectiveness of the DICE method drops, as the clock frequency increases.

The DICE method has lower hardware complexity comparing to LTMR one. So, it is preferable for use in the failure-resisted STSCs since they are inherently more complex than synchronous counterparts because of the dual-rail discipline and obligatory indication. In STSCs, LTMR method usage leads to a manifold complexity increase not only in the STSC's functional part but also in its indication subcircuit. As a result, not only the time parameters and energy consumption deteriorate, but also the reliability characteristics (for example, the average time between failures).

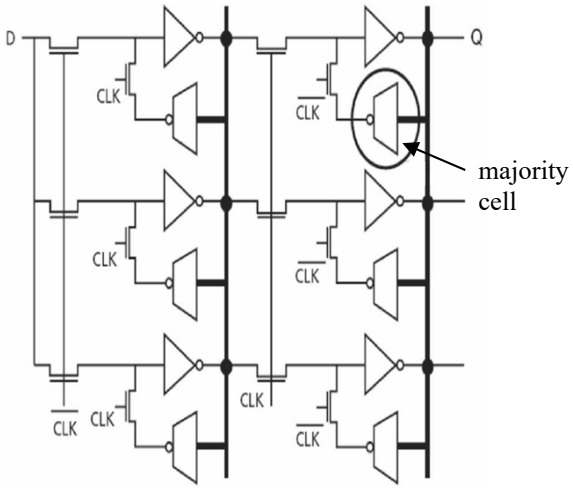
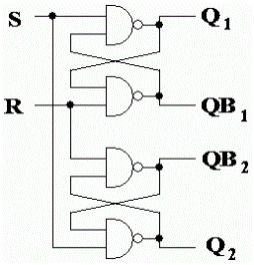


Figure 11 – Synchronous LTMR flip-flop

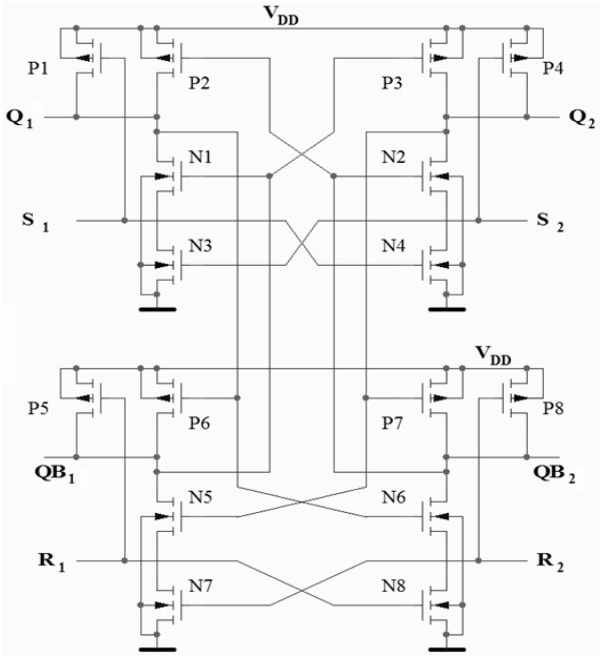
STSCs specificity makes it easy enough to adapt the DICE method for increasing their soft error tolerance. Figure 12(a) represents the synchronous simplified RS-latch circuit on NAND2 cells, which implements a DICE-like approach. The NAND2 cells are spaced a sufficient distance in a layout to minimize the probability of several particles impact them simultaneously.

Figure 12(b) shows a real DICE-style RS-latch circuit implementation [25]. Signals in the pairs $\{S_1, S_2\}$, and $\{R_1, R_2\}$ are logically equivalent and correspond to the S and R inputs in Figure 12(a). They are also spaced in the layout, as the increase in parasitic capacitance between them contributes to a growth in the sequential cell failure-tolerance [25]. Figure 12(b) demonstrates the real connections of the internal nodes of the circuit with the gates of transistors. Such the circuit successfully resists against the effects of a single particle in any submicron CMOS process.

However, an analysis shows that a significant difference in the time of switching its homogeneous inputs into the active state leads to the short-circuit current in the RS-latch. Let, for example, the RS-latch state in Figure 12(b) is $R_1=R_2=S_1=S_2=1$, $Q_1=Q_2=1$, $QB_1=QB_2=0$, and (R_1, R_2) inputs switch to logic-zero state ($R_1\rightarrow0$, $R_2\rightarrow0$), R_2 input does this later than R_1. Switching $R_1\rightarrow0$ results in unlocking the transistor P5, turning the node $QB_1\rightarrow1$, and unlocking the transistor N1. Since transistor N3 is already open ($S_2=1$), node QB_1 connects to the power supply through transistor P2 and to the ground through a chain of open transistors N1 and N3. Short-circuit current flows between the power and ground buses. It stops only after switching $R_2\rightarrow0$ and charging QB_2 node to a high potential, sufficient to close P2 transistor. Therefore, the R_1 and R_2, S_1, and S_2 nets should be paired with symmetry when implementing the RS-latch layout, so that the signal propagation delays along them are not too different.



(a)



(b)

Figure 12 – DICE-style RS-latch: schematic view (a) and CMOS circuit (b)

Using the DICE-like approach to implement STSCs also improves their failure-tolerance while complying with specific layout design rules. BSC built on the DICE-style RS-latch (Figure 12(a)) has two times more outputs than a conventional BSC. The DICE approach leads to doubling an STSC functional part. The principles of the ST circuits require an indication of all STSC inputs and outputs. Therefore, an increase in the number of outputs of the functional part also complicates the indication subcircuit.

Figure 13 shows a conventional ST RS-latch circuit, which information input is the DRS with a unit spacer. At the same time, the OAI22 cell implements an indication subcircuit. Figure 14 demonstrates the DICE implementation of the same latch. The indication subcircuit is more complex, as it should indicate all outputs in the ST circuit, but does not double.

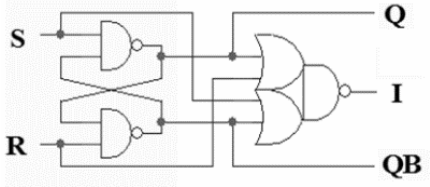


Figure 13 – Conventional ST RS-latch

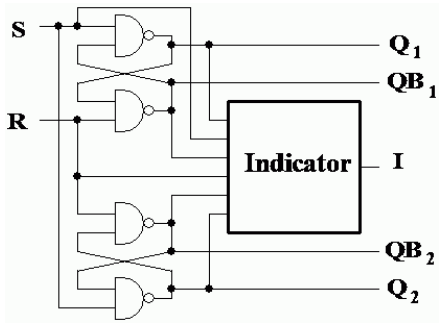


Figure 14 – DICE-style ST RS-latch

Figures 3 and 15 demonstrate a similar conversion of the original RS-flip-flop. The conventional ST RS-flip-flop shown in Figure 3 has two BSCs. In contrast, the failure-resisted DICE implementation shown in Figure 15 has duplicated BSCs. The indicator of the DICE implementation also is complicated but not doubled.

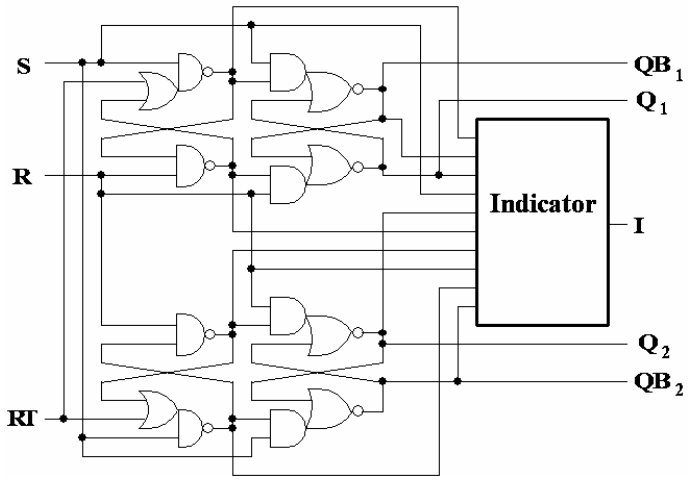


Figure 15 – DICE-style ST RSFF

Table 4: Data corruption probabilities in classical and advanced ST RSFF

Soft error case	Classical ST RSFF	Advanced ST RSFF	
		with DRS output	DICE-style
5.1	0.3125	0.3125	0
5.2	0.3125	0.3125	0
5.3	0.09375	0.09375	0
5.4	0.09375	0.09375	0
5.5	0	0	0
5.6	0.125	0.0625	0
5.7	0	0	0
5.8	0.125	0	0
5.9	0	0	0
5.10	0	0	0
Total	10.6%	8.8%	0%

Table 4 displays the data corruption probabilities in the ST circuit because of the soft errors, classified in Section 3 and analyzed in Section 5, for the ST RSFF implemented by the classical circuit (Figure 3), the circuit with DRS output (Figure 9), and DICE-style circuit. The last row shows the total data corruption probability because of soft errors in the RSFF

discussed in Section 5, taking into account the equal probabilities of their occurrence.

Thus, STSCs, unlike synchronous standard sequential cells and memory cells, mask 89.4% of soft errors analyzed in Section 5 due to their properties. The use of the ST RS-flip-flops with a DRS information output increases the tolerance to the soft errors discussed in Section 5, up to 91.2%. As in synchronous circuits, the DICE-style technique makes STSCs insensitive to remaining critical failures.

7. Conclusions

Due to using dual-rail code with spacer, two-phase operation discipline, and mandatory indication, ST circuits mask most of the short-term soft errors without any changes in their circuitry. Classical ST circuit implementations guarantee the high level of masking soft faults classified in Section 3: 91.4% in combinational ST circuits and 89.4% in ST sequential circuits.

The indication of prohibited DRS state that is opposite to the spacer state, as the second spacer masks remained classified soft errors in the combinational ST circuits.

The increase of the ST sequential circuits tolerance to the soft errors analyzed in Section 5 up to 91.2% is achieved by using ST RS-flip-flops with a dual-rail information output. DICE-style approach for implementing ST sequential circuits provides an additional increase of their tolerance to discussed soft errors up to 100%.

Thanks to ST circuit ability to mask most of the soft errors due to their features and their ability to maintain functional performance in a wide range of operating conditions, ST circuits are a promising basis for designing failure-resisted digital hardware.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

We would like to acknowledge the help of Dr. Alexander Kushnerov in discussing and preparing the paper. The study was funded by a grant from the Russian Science Foundation (Project № 19-11-00334).

References

[1] Y. A. Stepchenkov, A. N. Kamenskih, Y. G. Diachenko, Y. V. Rogdestvenski, D. Y. Diachenko, "Fault-Tolerance of Self-Timed Circuits" in 2019 10th International Conference on Dependable Systems, Services, and Technologies (DESSERT), Leeds, United Kingdom, 2019. https://doi.org/10.1109/DESSERT.2019.8770047

[2] A. I. Chumakov, "Forecasting local radiation effects in IC under the influence of the outer space factors" Microelectronics, 39(2), 85-90, 2010 (article in Russian with an abstract in English)

[3] P. Eaton, J. Benedetto, D. Mavis, K. Avery, M. Sibley, M. Gadlage, T. Turflinger, "Single event transient pulse width measurements using a variable temporal latch technique" IEEE Trans. Nucl. Sci., 51(6), 3365–3368, 2004. https://doi.org/10.1109/TNS.2004.840020

[4] B. Wiseman, "Design and Testing of SEU/SEL Immune Memory and Logic Circuits in a Commercial CMOS Process" in IEEE Radiation Effects Data Workshop, 1994, 51-55. https://doi.org/10.1109/REDW.1994.633039

[5] S. I. Ol'chev, V. Y. Stenin, "CMOS logic elements with increased failure resistance to single-event upsets" Russian Microelectronics, 40(3), 156-169, 2011. https://doi.org/10.1134/S106373971103005X (article in Russian with an abstract in English)

[6] J. E. Knudsen, L. T. Clark, "An area and power-efficient radiation hardened by design flip-flop" IEEE Trans. on Nucl. Sci., 53(6), 3392-3399, 2006.https://doi.org/10.1109/TNS.2006.886199

[7] Y. Monnet, M. Renaudin, R. Leveugle, "Hardening techniques against transient faults for asynchronous circuits" in 11th IEEE International Conference: On-Line Testing Symposium, IOLTS 2005. https://doi.org/10.1109/IOLTS.2005.30

[8] S. F. Tyurin, A. N. Kamenskih, "Research into the reservation of logic functions at transistor level" In a world of scientific discoveries, 10(58), 232-247, 2014. https://doi.org/10.12731/wsd-2014-10-18 (article in Russian with an abstract in English).

[9] C. LaFrieda, R. Manohar, "Fault Detection and Isolation Techniques for Quasi Delay-Insensitive Circuits" in International Conference on Dependable Systems and Networks, 2004. https://doi.org/10.1109/DSN.2004.1311875

[10] A. Balasubramanian, B. L. Bhuva, J. D. Black, L. W. Massengill, "RHBD techniques for mitigating effects of single-event hits using guard-gates" IEEE Trans. on Nucl. Sci., 52(6), 2531-2535, 2005. https://doi.org/10.1109/TNS.2005.860719

[11] M. T. Moreira, G. Trojan, F. G. Moraes, N. L. V. Calazans, "Spatially distributed dual-spacer null convention logic design" Journal of Low Power Electronics, 10(3), 313-320, 2014. https://doi.org/10.1166/jolpe.2014.1332

[12] S. A. Aketiy, S. G. H. Cheng, J. Mekiey, P. A. Beerel, "SERAD: Soft Error Resilient Asynchronous Design using a Bundled Data Protocol," arXiv:2001.04039 [cs.AR]. http://https://arxiv.org/abs/2001.04039v1

[13] A. Taubin, A. Kondratyev, J. Cortadella, L. Lavagno, "Behavioral Transformations to Increase Noise Immunity in Asynchronous Specifications," in International Symposium on Advanced Research in Asynchronous Circuits and Systems, 1999. https://doi.org/10.1109/ASYNC.1999.761521

[14] D. Muller, W. Bartky, "A theory of asynchronous circuits" Annals of computation laboratory of Harvard University, V.29, 204-243, 1959.

[15] K. M. Fant, S. A. Brandt, "NULL Convention Logic: a Complete and Consistent Logic for Asynchronous Digital Circuit Synthesis" in International Conference on Application-Specific Systems, Architectures, and Processors, 1996. https://doi.org/10.1109/ASAP.1996.542821

[16] M. Kishinevsky, A. Kondratyev, A. Taubin, V. Varshavsky, Concurrent Hardware: The Theory and Practice of Self-timed Design, J. Wiley, 1994.

[17] Y. A. Stepchenkov, Y. G. Diachenko, V. S. Petrukhin, A. V. Filin, "The price of realizing the unique properties of self-timed circuits" Systems and means of informatics, 9, 261-292, 1999. (article in Russian with an abstract in English)

[18] Y. A. Stepchenkov, Y. G. Diachenko, G. A. Gorelkin, "Self-Timed circuits are the Microelectronics Future" Radio-electronic issue, 2, 153-184, 2011 (article in Russian with an abstract in English).

[19] Ю. А. Степченков, А. Н. Денисов, Ю. Г. Дьяченко, Ф. И. Гринфельд, О. П. Филимоненко, Н. В. Морозов, Д. Ю. Степченков, Л. П. Плеханов, Библиотека функциональных ячеек для проектирования самосинхронных полузаказных БМК микросхем серий 5503/5507, Техносфера, 2017. ISBN 978-5-94836-332-5 (in Russian).

[20] D. Shang, A. Yakovlev, A. Koelmans, D. Sokolov, A. Bystrov, "Registers for Phase Difference Based Logic" IEEE Trans. Very Large Scale Integration (VLSI) Systems, 15(6), 720-724, 2007. https://doi.org/10.1109/TVLSI.2007.898772

[21] Y.-C. Rhee, "Exclusive-OR and/or Exclusive-NOR Circuits Including Output Switches and Related Methods," US Patent No. 7,312,634 B2, 2007

[22] Y. Stepchenkov, Y. Rogdestvenski, Y. Diachenko, D. Stepchenkov, Y. Shikunov, "Energy Efficient Speed-Independent 64-bit Fused Multiply-Add Unit" in 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus2019), 2019. https://doi.org/10.1109/EIConRus.2019.8657207

[23] M. Berg, "Revisiting Dual Interlocked Storage Cell (DICE) Single Event Upset (SEU) Sensitivity" in Microelectronics Reliability & Qualification Work Meeting (MRQW) 2013 and HiREV Industry Day, El Segundo, CA, December 10-12, 2013.

[24] T. Lakshmivaraprasad, M. Sivakumar, B. K. V. Prasad, S. A. Inthiyaz, "Nanoscale CMOS technology for hardened latch with efficient design" International Journal of Electronics and Communication Engineering, 5(3), 343-349, 2012. ISSN 0974-2166

[25] Y. V. Katunin, V. Y. Stenin, P. V. Stepanov, "Simulation of trigger two-phase CMOS logic cell characteristics, taking into account the separation of charge at the effects of individual nuclear particles" Microelectronics, 43(2), 104-117, 2014. https://doi.org/10.7868/S0544126914020069 (article in Russian with an abstract in English)

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 57-63 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Performance Evaluation and Examination of Data Transfer: A Review of Parallel Migration in Cloud Storage

Mimouna Alkhonaini*, Hoda El-Sayed

Department of Computer Science, Bowie State University, 20715, USA

A R T I C L E I N F O

Article history:
Received: 05 November, 2019
Accepted: 21 February, 2020
Online: 09 March, 2020

Keywords:
Parallel
Transfer
Cloud

A B S T R A C T

Recent years have seen a continued pattern of development in the cloud computing field. Numerous approaches to maximize file transfer capacity are still completely standing for use on cloud computing storage; however, they do not maximize the advantage of data migration scalability and elasticity in cloud storage. One potential problem is that elasticity takes time; however, the scalability attributes that have not been fully exploited include multicore chips and parallelization that can further be leveraged to enhance the overall data transfer performance and efficiency. In that regard, considerable effort has been directed to multiprocessors. Such systems involve a plurality of processors or functioning units capable of independent operation to process separate tasks in parallel. Nevertheless, the penalization is complicated when a task requires several resources or signals to proceed with meaningful computation. Thus, accommodating equitable priority among tasks further complicates operations. In this paper, we propose a parallel server to cloud storage transfer system in which parallelism method can only be utilized in case of transferring a large number of files and applied in order to increase the transfer throughput. The data is transmitted into several chunks via TCP network within the same period slot in a single data path which indicates dataflow on parallelism. Our target in this system is that increasing number of processors and the problem size will simultaneously maintain the efficiency of the data transfer system. The proposed model is based on the combination of dynamic segmentation, CRS, AES, and hashing. In summary, the proposed model shows the potential to enhance the performance by increasing the data transferability. The performance of the proposed model will be measured with the help of comparing the average execution time with the number of processors and speedup of the entire parallel system.

1. Introduction

The cloud computing involves the provision of on-demand computing services to the customers. These services are similar to the usual services that can be obtained by using the physical computing equipment such as storage devices, servers, networking and many more. As the realization of the benefits of cloud computing come to light, many businesses, companies, organizations, and in stitutions are switching to the cloud services. Cloud services are provided by the cloud service providers (CSP) who are responsible for manning all the infrastructures, software, and the platforms to which the customers subscribe. Performance remains a significant concern for the consumers who are still struggling with the decision of moving to the cloud. Though there are many significant standards in cloud computing, the primary standard is on the data migration performance from the non-cloud infrastructure to the cloud. This paper is an extension of the work which was initially published and presented in IEEE 20th International Conference on High-Performance Computing and Communications [1].

*Mimouna Alkhonaini, COSC Department, 14000 Jericho Park Road, Bowie, MD 20715, 301-860-3964 & alkhonainim0522@student.bowiestate.edu

Typically, a multiprocessor includes a variety of computational units, a memory, a control, and at least one I/O processor. The functions of a job are initiated and processed, sometimes being moved from one computational unit to another. In the course of such operations, tasks must be synchronized and, in this regard, a function may require data from another function supplied through the I/O processor [2]. Under such circumstances, there is an issue in increasing number of processors against problem size.

The features of parallelism have been known since Babbage's effort to construct a mechanistic computer [3]. Industrial and academic researchers have studied every imaginable aspect of parallel computation. A system becomes slower due to combined cumbersome computation processes, therefore, there is a need of using parallel computing to improve and provide high levels of execution performance efficiency. Moreover, this research also provides the design and validation of the analytic performance of the parallel migration to enhance the speed of the system.

Related studies in the fields of data migration via parallelism have gained focus over the past few years. However, most of the proposed studies only get the standpoint of either a cloud provider or cloud consumers, not both. In the present work, the proposed scheme shall concentrate on allowing the cloud provider and also the consumer to accomplish data migration efficiently by exploiting potential improvements found in consumer PC multicore chips. Moreover, future cloud storage environments will be primarily designed and built on top of multicore technology.

In this paper, we propose a parallel migration system that can be made cost-optimal by adjusting the number of processors and the problem size. The parallel migration is achieved by dynamic distribution for the combination of file slicing, cryptography, fault tolerance, and hashing.

The remaining paper is structured in a manner that section 2 presents the related works in this field. Section 3 defines and models how the proposed system achieves parallel migration by dynamic distribution. Section 4 shows the experimental outcome and section 5 concludes the paper with some future remarks.

2. Related Works

In this section, we discuss some of the approaches related to the data migration parallelism as well as increasing file transfer bandwidth.

HTTP, SFTP, and SCP are the standard data migration communication protocols for point-to-point transfer. In terms of elasticity, the improvement is based on the work of [4] that helps to scale down the repetitive data migration through developing "checksum comparison" and the work presented in [5] prevents resource conflict by distinguishing "back-pressure". However, still, these approaches have thus far to capitalize on cloud elasticity. In our method, we aim at matching the amount of resources allocated to data transfer in parallel mode with the amount of resources it actually requires, avoiding over-provisioning or under-provisioning.

In terms of connection speed, the existing odd source to odd end target data migration speed is usually limited by either cloud provider settings or limits set by the hardware of physical connection. Though, one of the solutions to prevent these restrictions and to expand data migration throughput is to perform data migration from multiple servers concurrently. Even though, BTorrent [6] illustrates that this is rational but processing this data involvement procedure (As It Is) is far from appropriate "for a one-time point-to-point" migration. However, our parallelization data migration combines data segmentation and erasure code, AES, and hash with developing parallel migration capacity and minimizing data transfer delay.

Multiple data transmission develops the processor's total bandwidth; however, it is limited because the "bottleneck" lies in the specific cloud processor's bandwidth. One more comparable system is the parallel transfer where files are prearranged through several input/output processors to minimize storage bottleneck, which allows transmission across several TCP connections concurrently. Another similar system can be applied in cloud storage, but the overheads to rearranged files across numerous storages should to be considered in order to improve the file transmission performance. In our parallelization model, provisional buffer is created on the cloud machine, and these buffers are indicted to as the 'holes'. As the buffer hole is fully occupied, it is then taken away from the list. The entries of theses buffer-holes are then tested with the descriptor-list to confirm if the buffer-hole which has been loaded by the arriving segment is rejected [1].

In [7], multi-hop and multi-path which are two optimization techniques to increase the performance of file transmission across WAN are briefly discovered. In multi-hop path splitting, throughput was improved by exchanging direct connection between start and end point via multi-hopping. On other hand, multi-path technique includes slicing data at the start point and sending it through various overlapping paths. Chunking and rebuilding data theoretically cause data corruption, damage or failure. Thus, our proposed model takes into consideration fault-protection by applying several Reed-Solomon (RS) code values along with the data sizes to supplementary explore the impact of different values.

The authors in [8] propose a cloud-based data management system targeted for big data science applications running across large and high geographical environment sites which shows expectable data transfer management performance in terms of cost and time. The system automatically implements and adjusts performance designs for the cloud platform structure for enhancing data and efficient schedule of the transfer process. Though, it shows a poor execution performance when compared with our proposed parallelization model.

Another similar approach [9] is the parallel data transfer where data are lined across several intermediate nodes that are spawned in order to avoid the bandwidth limitation, but the overheads to reallocated data across multiple nodes in source DC and destination DC have to be taken into attention for performance improvement on the file transfer.

3. Server to Cloud Parallel Data Transfer Modelling

3.1. General Concept of the Model

In this section, we define the outline of our server-to-cloud parallel transfer conception and formulate an arithmetic model constructed on the proposed parallel migration concept. In the rest of the paper, we then cover the server-to-cloud parallel migration task as parallel data transfer.

Figure 1 represents the construction overview of our parallel migration model. The process starts with the data as the input, the data is then segmented into chunks, every segmented data is

encrypted. Encrypted chunks are migrated to the cloud and then decrypted. The decrypted chunks are then compared against the un-encrypted segments on the local server. If they don't match, the process is taken back to the segmentation stage. If they match, the segments are then re-assembled and saved on the cloud server. We then end the process.

The following is the arrangement of status for a single file transmission from server to cloud storage.

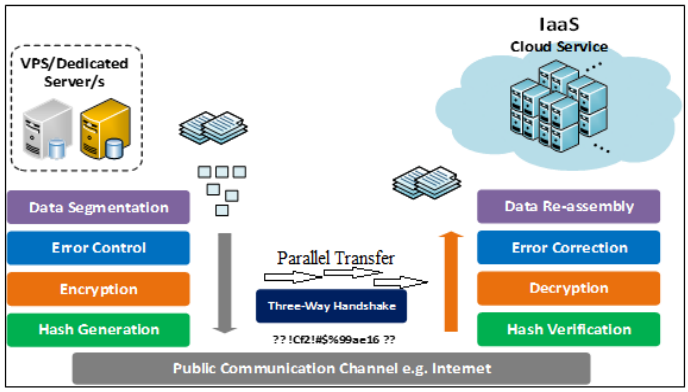


Figure 1: Overview of parallel transfer mechanism

- Data Segmentation [10]: Data segmentations split-up the bigger data units into smaller controllable data slices. In our data segmentation, we will be using an algorithm that will take the shortest time possible to segment larger volumes of data packets, while optimizing the memory efficiency of the working machines. In this algorithm, the system scans the data packets to find the size of the whole data to be segmented. The total data size is then used as a base for the determination of how many segments should be generated from the original data. Data size levels are then predefined to allow the categorization how a given data set belonging to a given size group is segmented. The categorizing of data based on the whole data size ensures that we have larger data size, the right segmentation size will be obtained in order to make the process faster. The data will then be split into smaller equal sized chunks of 4MB each. The chunks are assigned a sequence number each, according to the order in which they are segmented. If we have a larger volume of data, then more instances of segmentations will be initiated to make sure that less time is taken to process the whole data. The number of segmentation instances is dictated by the size of whole data.
- Error control: The end-to-end transmission of data from the server to the destination storage encompasses many steps which can be subjected to errors leading to data loss. Error control ensures that any error that occurs during the transmission of the data is detected and addressed. Reliable data transmission implements error detection and correction. Error correction enables a system to detect and correct both the bit-level and packet-level errors. There are two types of errors that can occur during the data transmission: single-bit error, in which only one-bit changes from 0 to 1 and vice versa and burst error where two or more bits change from 0 to 1 and vice versa. Burst-error is known as the packet-level error which results in duplication or re-ordering of the data. We are using the checksum method for error control which is applied at the upper-layer, and it uses the concept of redundancy. There are two operations which are being performed in checksum technique: Checksum-generator on the server side and Checksum-checker on the cloud side while the local server makes use of the checksum-generator. The following is the procedure for the checksum generation:
  - The data segments are sub-segmented into equal n-bits.
  - This is followed by adding together the sub-segments with the usage of 1's compliment.
  - The data segment is re-complimented to become a checksum.
  - It is then sent alongside the data-unit.
- Encryption: The system uses an Advanced Encryption System (AES) with the RSA key-exchange. The symmetric encryption technique requires a proper key exchange to maintain the security of communication and minimize the risk of eavesdropping during the key exchange phase. RSA-2048 is then used for key exchange. The main reason for not choosing RSA for encryption is that it needs more processing power and is slower than the mighty AES.
- Hash generation: Hash functions are one-way which means that it is impossible to revert from the generated values to the initial values. The hash function is used in our system to enhance the confidentiality of the data (message). We are using the secure hashing algorithm (SHA) which was modeled after MD4 and its proposal was made by NIST to offer a more secure hash standard (SHS). This will produce a 160 bits hash value.
- Three-way handshake initialization: Three-way handshake is a flow control technique which allows the communicating devices to initiate the communication platform. In Figure 2, the local server will initiate the three-way handshake by sending the cloud server a SYN (synchronization) message. This is a way of asking the cloud server if it is ready to communicate. The recipient, which is the cloud server sends back a SYN/ACK message to the local server telling it that it is ready for the communication. The local server then sends an ACK (Acknowledgement), and the connection between the two sides is established. The data is then transmitted.

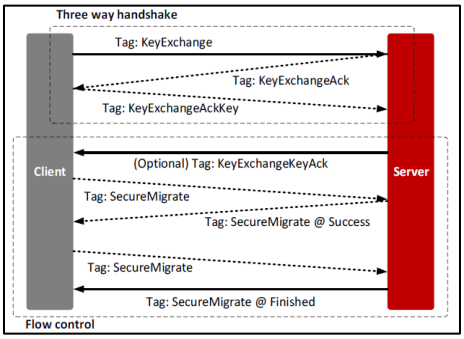


Figure 2: Three-way handshake initialization.

- Hash verification: Hash checks for the integrity of the message. The cloud server receives a segment and runs SHA algorithm over it, if it matches, then we increment the number of received healthy segments by one, we ignore (discard) the rest of the segments which do not match. The system waits until receiving other segments (we know the total number of segments sender

is going to send for us) and see how many of them are healthy (their integrity can be verified).

- Decryption: With the Cipher Block Chaining decryption, application of the inverse-cipher function to the first cipher-text block results into an output-block which passes through the XOR operation with the initialization-vector to allow the recovery of the first plaintext block. The inverse-cipher function is also applicable to the second cipher-text block which provides the output-block which is then XORed with the first cipher-text block for the recovery of the second plaintext block. Generally, the recovery of any plaintext block (except the first plain-text), the inverse cipher-function applies to the corresponding cipher-text block, and the output-block is XORed with the prior cipher-text block.
- Error correction: Error correction allows the cloud server which receives the data to reconstruct the original data in the event that it was corrupted during transmission. Error correction is implemented using the hamming code. Hamming code is a single-bit error correction method that uses the redundant bits. The redundant bits are included on the original data and their arrangement is in such a way that any different and incorrect bits will lead to the production of various error results. The corrupted bits are then identified and upon the identification of the corrupted bits, the cloud server which receives the data is capable of reversing its values hence, correcting the errors. Hamming code applies to any data length and makes use of the relationship between the data and the redundancy-bits. Checksum-checker: The cloud server makes use of the checksum-checker. The following is the procedure for the checksum checker:
    - The cloud server which receives the data-unit will divide the received data-unit into equal sub-segments.
    - All the sub-segments are then added together using the 1's compliment.
    - The outcome is complimented again. If the final result is found to be zero, the system accepts the data while when it is found to be 1, the data should be rejected and considered to have errors.
- Reassembling of data [11]: After all the data segments have arrived and each segment is verified, an algorithm to arrange the segments in order according to their sequence numbers will be used. This ensures that the data segments will be ordered back to their original state before being segmented on the local server [12]. In Figure 3, the flowchart diagram shows how the model receives the segment and allocates the buffer space for each segment. There are more segments in the buffer, by comparing the size of the segment with the size of the space available in the buffer, the segment is taken back. If there are no more data in the buffer, then the data unit is identified, and the data unit is bound to the context. It will then check if it is occupied by a different data unit. The bound data is then updated, and the segment is inserted into the data unit. The number of segments is then checked, if there are any remaining segments, they will return back to step one. If there are no remaining segments, then the header will be updated and the data unit will be transmitted. The numbering sequence plays a crucial role as it is the sequence number used to arrange back the data to its original state [13]. The size of the unencrypted segments on the source server is furthermore tested and matched with the size of the corresponding decrypted segments on the destination server. Re-assembling of data segments will only take place after all the previous stages have been successfully completed.

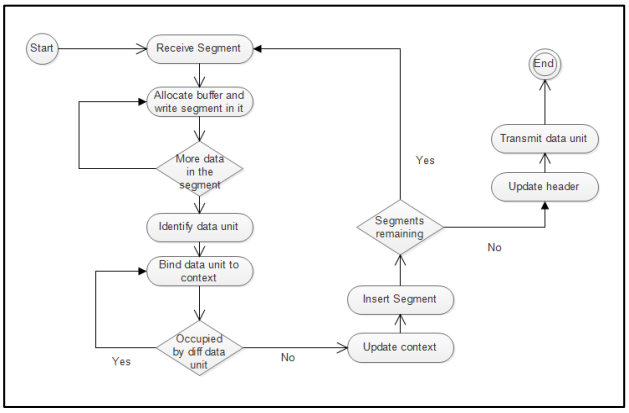


Figure 3:Reassembling of segments.

3.2. Parallel Transfer Algorithm

Server user may perhaps definitely benefit from the rapid growth of multicore systems. The physical server typically holds several regular data centers or computer storage, where commodity high-performance multicore bunches are employed. Our migration algorithm designed for one direction data flow which begins at the server storage and ends at the cloud storage. Algorithm 1 shows the parallel-serial transmission which starts from a master processor acting as a management system. The aim of the master processor is to reach the final output by collecting the necessary information from the slave processors. In our migration algorithm, we will assume that the data is ready for transfer after the four stages: data chunks, error control/correction, encryption/ decryption with AES-256 and SHA-256 algorithm which is identified as the significant component of the integrity process. The transfer process will start by taking the ready data list, FList, and divides it between the slave processors. Correspondingly, each slave processor would obtain blocks of FList. We use single program multiple data (SPMD) algorithm where there is a single instruction that each processor performs, but on various data. Each slave processor would establish its secure channel, port, and buffer. Each slave processor will then send its Flist to the cloud buffer before the storage stage. The final task of the slave processor is to update its Flist by removing the sent chunks from the list. This procedure would reflect to have a higher bandwidth transmission capacity. Finally, the master processor receives the FList from all the slave processors and takes the re-transmission decision. If the Flist size=0, it means that all the data chunks have been sent successfully. However, if Flist size=!0, it means that the master processor will request for the re-transmission of the missing chunks by the method explained in [1]. The master processor would close the connection after all the re-transmission requests. At this phase, the cloud buffer is full of the data that need to be checked for integrity, verified and re-build to be stored.

3.3. Time Factor of Parallel Transfer

The performance of the proposed model will be evaluated by computing the average of execution time, the number of processors, and the speedup ratio.

Algorithm 1 Parallel Migration Algorithm

Input: FinalChunksList FList

```
procedure AllocatesBlockToEachProcessor
    P ← Number of processors
    if Processor 0 then                    ▷ Processor 0 is the root
        Scan FList size
        Establish Secure Connection (3-way handshake )
        Divide the FList into x by y blocks        ▷ Matrix
        for i ← 0, x − 1 do
            for j ← 0, y − 1 do
                Send block (i, j) to processor (i+j) mod P
            end for
        end for
    end if
end procedure
procedure EachProcessorTransmission
    (Note that RS,AES25,SHA5126 is applied on FList)
    Receive FList blocks from the root
    S ← Scan block size
    for index ← 0, S − 1 do
        Initiate the TCP
        Transfer the block
        Remove the block (FList[index])
    end for
    Returen FList
end procedure
if Processor 0 then
    if FList Non Empty then
        while NotAllFListsSent do
            send the next FList
        end while
    else Close Connections
    end if
end if
```

3.3.1 Data Distribution Time (Segmentation+ CRS+ AES + HashGen)

Divided into four actions, file chunking, encoding, encryption, and hash generation. In order to minimize the delay between individual execution time, transferring has to be started directly after the segmentation when the exact chunk is ready, without waiting for the overall segments process to be completed. Server execution time can also be minimized by manipulating the chunk size and sequence when the network resource is unutilized.

3.3.2 Parallel Data Migration Time

This time factor indicates the transfer from local server to buffers in the cloud storage. In order to increase the throughput of this phase, the parallelism method can be applied only if the transmission files are of a large number. Processor 0 would act as the root and starts to define the number of the chunks and divide them among the slave processors. As a result, each slave processor would collect a block of chunks. Later, they would establish their connections and send their chunk lists to the cloud buffers. The control would then be passed to TCP in order to finalize the transmission task.

3.3.3 Data Amalgamation Time (HashVerify +AES +CRS +Re-Assembling)

In this time factor, the file reassembling is the prime task. It also involves the migration from the cloud buffers to the destination storages. This step also contains massive disk operations such as the writing process.

3.4. Model and Evaluation of Parallel Migration

The parallel performance was analyzed by displaying speedup rate and throughput to illustrate the parallel model profit. The speedup of parallel operations on P processors is defined as the following [14]:

$$S_P = \frac{T_1}{T_P} \tag{1}$$

Where T_1 is the sequential execution time using one processor, T_P is the parallel execution time using P processors, p and 1 are indexes as written in equation 1. While Sequential Transfer Time equation (T_1) is:

$$T_1 = \frac{D_s}{t_e} \tag{2}$$

Where D_s is data size in MB, t_e is transfer speed in MB/s. Parallel Transfer Time (T_P) = data distribution time in the server + transfer time + data amalgamation time in the cloud. Hence, the speedup rate is acquired as below:

$$S_P = \frac{D_s/t_e}{\frac{T_P}{D_s/t_e + D_s(1/P)/t_e}}$$

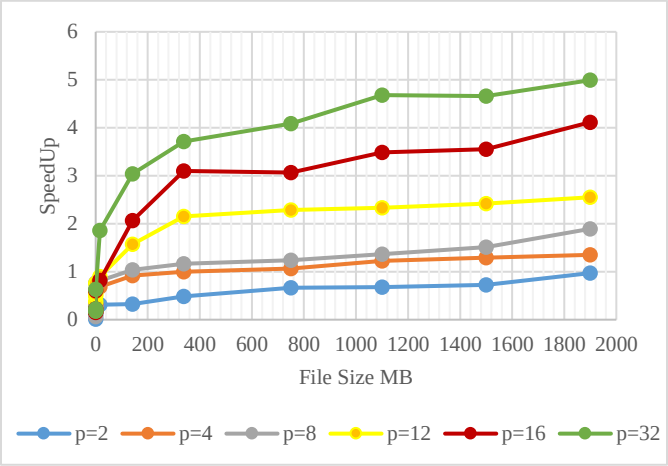


Figure 4: Speedup model for parallel transfer, where p is the number of processors.

It can be noticed that the smaller file transfer size (not many of small and large files together), a minimum number of parallel processors are required to accomplish a speedup. However, when trying to parallelize the transfer, there is no prediction of when the packet loss rate would begin to increase exponentially which causes a critical overhead. Thus, smaller file sizes do not benefit from the parallelization. Therefore, Figure 4 shows that the speedup rate is steep for the file size of below 4MB.

For a larger data transfer size, the speedup rate grows as the number of parallel processors are increased. However, performance gains are hardly worth when each subsequent increase in processors number brings a lesser performance profit contrasted to the previous increment.

4. Experimental Evaluation

This section shows the performance evaluation for our parallel migration system on an Infrastructure as a Service (IaaS) cloud environment.

4.1. Experiment Setup

We carry out these experiments on two separate VMs (VM hardware were identical) hosted on Cray XC40 within two locations: Oregon and Amsterdam. Linux Operating System in various cores with 8 GB of memory was assembled on the VMs. All the encryption /decryption stages are constructed with the help of python cryptography 2.1.4.

Our experiment concentrates on migrating a single file from the server to cloud storage. The model is designed in such a way that each part of the system is pluggable. Thus, replacing an encryption algorithm will not affect the selection of the coding algorithm. We have chosen numerous file sizes, starting from .05 MB to 1900MB and RS parameters (k, m) in range of ($1 \leq m \leq k$ and $k \leq 256$, where m=parity) in order to evaluate the transfer rate. For coding layer, we employed PyEClib source on GitHub (1.5.0) module with liberasurecode source on GitHub (1.5.0) as Interface to erasure back-ends for Reed-Solomon coding.

4.2. Preliminary test

To evaluate the actual performance, we first examined the performance of our algorithm by measuring the execution time in the server (source) and cloud storage (destination) processes for a data, depending on the file size and variability of RS parameters k and m.

In Table 1 below, we showed the average operating time for the source and destination for each process. It is quite apparent that the source average is mainly estimated by the slicing process because it involves additional parameters and computations than the other internal operations.

Table 1: Average execution time in the source/destination.

OPERATIONS	TIME (ms)	SUM
CRS Coding	691.99	1,089.74
AES Encryption	116.78	
SHA Hash	280.97	
SHA Verify	277.39	487.55
AES Decryption	119.53	
CRS Decoding	90.63	

The throughput define as that is the amount sum of processed data that execute in a period of time (processed data /second) [15]. It is concluded as the following relation:

$$P_n = \frac{W}{T_P} \qquad (3)$$

Where W is indicated as the workload, T_P as the execution time required to accomplish the calculation using P processors [16], p is index as written in equation (3). From equation 3, the throughput for each location is shown in the Table 2. Table 2 depicts the comparison between the average throughputs of transmitting many file sizes between Oregon and Amsterdam locations. The transfer throughput gained for Oregon and Amsterdam is ~.9, .7 MB/ms respectively. The range of throughput achieved when transferring the file is astonishingly large. It was observed that the throughput rate does not follow a certain way; the variation seems to arise randomly.

Table 2: Locations Transfer Throughput.

File sizes	Oregon TO Amsterdam	Amsterdam TO Oregon
0.05	0.005101482	0.001007
0.10	0.009208418	0.001696
0.55	0.038767525	0.008251
1.00	0.019112421	0.011328
17.00	0.041170096	0.073851
142.00	0.086551034	0.263185
339.00	0.093888415	0.333905
750.00	0.107918311	0.442913
1,100.00	0.102385909	0.468758
1,500.00	0.109188015	0.502546
1,900.00	0.114588619	0.579061

4.3. Result

Figure 5 and Figure 6 are linearly dependent features and are just shown for further clarification. It is verifiable that by choosing bigger fault tolerance rate in the coding stage (smaller k in proportion to m) the final transfer time increases rapidly. It is shown that transferring bigger files with no-error tolerance (k =m) added in the coding stage, the system will exhibit the best performance, especially in case of bigger file sizes. Choosing k such that it covers 10-15% error correction without the need for data retransmission in smaller file sizes less than 24MB, the impact on transmission time is noticeably less than the profit it brings.

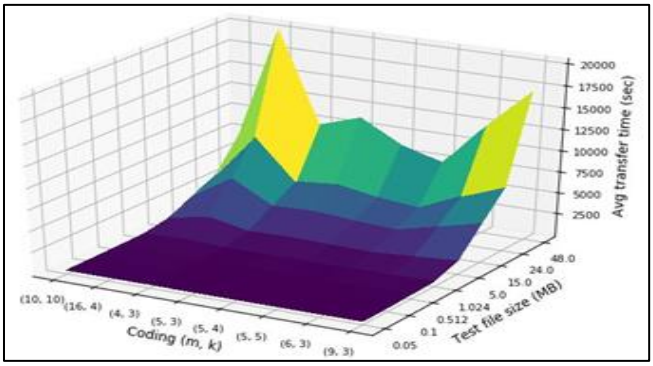


Figure 5: Average execution time in source side.

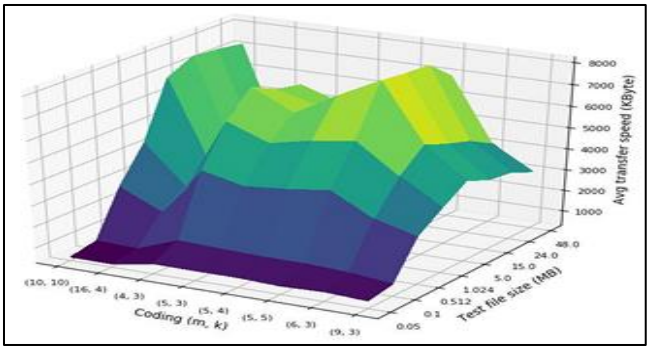


Figure 6: Average execution time in destination.

In Figure 7, we portray a 2D graph of the speedup rate against data size for a different number of processors. It can be noted that analogous to our model, speedup cannot be obtained for the data migration that is below 0.5GB by using 2, 4, 8, and 16 parallel processors. The overall note is steady with our approach, the speedup is extremely poor for small file size, but increases for bigger file size.

Although the speedup rate is not computed yet in the real environment, we found an independent argument that can be used in our model to compute the speedup in the real environment. Based on our observation, the independent argument is close to 0.47. Hence, we are certain about the fact that when the count of parallel processors is greater than 8 and the data size is bigger than 1GB, the speedup will be achieved.

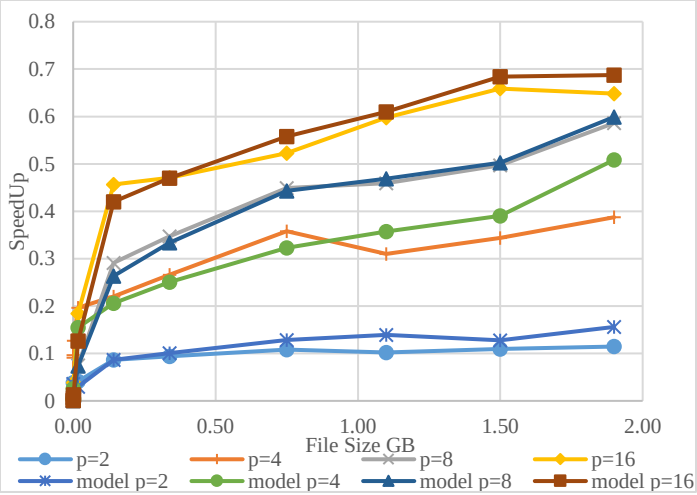


Figure 7: Comparing actual speedup and our model for parallel transfer, where p is the number of processors.

5. Conclusion and Future Work

This paper concentrated on explaining and exhibiting a parallel design for data transmission in cloud environment. First, we demonstrated the proposed model. Next, we presented the implementation and evaluation of the proposed model. Though our parallel transfer model shows a prospect to enhance the performance, however, our implementation exposed that there is an independent argument that should be considered. This study also suggests that there are overhead factors that can affect the parallel transfer performance, such as initialization, chunking the data file, and transfer time. Thus, those overhead effects should be investigated in order to reach maximum performance. In brief, we should consider the computation of hardware power and processors I/O's speed that is assigned by the CSP to design the parallel migration model.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] M. Alkhonaini and H. El-Sayed, "Optimizing Performance in Migrating Data between Non-cloud Infrastructure and Cloud Using Parallel Computing," *2018 IEEE 20th International Conference on High Performance Computing and Communications (HPCC)*, Exeter, United Kingdom, 2018, pp. 725-732.

[2] Hesham El-Rewini, Hesham H. Ali, and Ted Lewis. 1995. Task Scheduling in Multiprocessing Systems. *Computer* 28, 12 (December 1995), 27-37. DOI: https://doi.org/10.1109/2.476197

[3] S. Razdan, Fundamentals of Parallel Computing. New Delhi: Alpha Science International, 2014.

[4] "rsync(1) - Linux man page." [Online]. http://linux.die.net/man/1/rsync. 2015.

[5] H. Pucha, M. Kaminsky, D. G. Andersen, and M. A. Kozuch, "Adaptive File Transfers for Diverse Environments," in USENIX 2008 Annual Technical Conference on Annual Technical Conference, Berkeley, CA, USA, 2008, pp. 157–170.

[6] Bram Cohen, "The BitTorrent Protocol Specification." [Online]. http://www.bittorrent.org/beps/bep_0003.html. 2015.

[7] G. Khanna et al., "Multi-hop Path Splitting and Multi-pathing Optimizations for Data Transfers over Shared Wide-area Networks Using gridFTP," in Proceedings of the 17th International Symposium on High Performance Distributed Computing, New York, NY, USA, 2008, pp. 225–226.

[8] R. Tudoran, A. Costan, R. Wang, L. Bouge, and G. Antoniu, "Bridging Data in the Clouds: An Environment-Aware System for Geographically Distributed Data Transfers," in Cluster, Cloud and Grid Computing (CCGrid), 2014 14th IEEE/ACM International Symposium on, 2014, pp. 92–101.

[9] C. B. M. Lek, O. B. Yaik and L. S. Yue, "Cloud-to-cloud parallel data transfer via spawning intermediate nodes," TENCON 2017 - 2017 IEEE Region 10 Conference, Penang, 2017, pp. 657-661. doi: 10.1109/TENCON.2017.8227943

[10] M. Alkhonaini and H. El-Sayed, "Minimizing Delay Recovery in Migrating Data between Physical Server and Cloud Computing Using Reed-Solomon Code," *2018 IEEE 20th International Conference on High Performance Computing and Communications (HPCC)*, Exeter, United Kingdom, 2018, pp. 718-724.

[11] M. Alkhonaini and H. El-Sayed, "Migrating Data Between Physical Server and Cloud: Improving Accuracy and Data Integrity," *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, New York, NY, 2018, pp. 1570-1574.
doi: 10.1109/TrustCom/BigDataSE.2018.00226.

[12] "Sequence data submission and accession numbers", FEBS Letters, vol. 360, no. 3, pp. 322-322, 1995.

[13] C. Diot and F. Gagnon, "Impact of out-of-sequence processing on the performance of data transmission", Computer Networks, vol. 31, no. 5, pp. 475-492, 1999.

[14] Roosta, S. " Parallel Processing and Parallel Algorithms: Theory and Computation ", ISBN: 0-387-98716-9, Springer Verlag, 2000.

[15] Kermarrec, A.; Bougé, L.and Priol, T. "Euro-Par 2007 Parallel Processing ", 13th International Euro-Par Conference: Lecture Notes in Computer Science, ISBN 978-3-540-74465-8, Vol. 4641, 2007.

[16] Hwang, K. and Xu, Z. (1998). Scalable Parallel computing. New York: McGraw-Hill.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 64-73 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Regenerative Energy Recovery System for Electric Vehicles Charging A Battery at A Low Speed

Wonseok Yeo, Sungchul Jung, Seungtaik Kim, Keonho Park, Jongsun Ko*

Dankook University, Electronics and Electrical Engineering, Institute, KS009, Republic of Korea

ARTICLE INFO

Article history:
Received: 04 December, 2019
Accepted: 01 February, 2020
Online: 09 March, 2020

Keywords:
Regenerative Energy
Energy Recovery
Battery
Electric Vehicle
L-C Resonance
Permanent magnet synchronous motor

ABSTRACT

In this paper, a regenerative energy recovery system for electric vehicles charging a battery at a low speed is proposed. When a permanent magnet synchronous motor driving the electric vehicles is driven by a generator, the generated voltage is controlled to rise or drop by using the L-C resonance circuit. As the circuit only has three capacitors and three IGBTs, the proposed system is very simple. A resonance frequency is found through frequency analysis of the L-C resonance circuit; furthermore, using the resonance effect, a generated voltage control algorithm is proposed. PI duty control in the algorithm raises the generated voltage and reverse PI duty control in the algorithm reduces the generated voltage. The results of simulations and experiments confirm that a remodeled electric vehicle by the proposed system is charged to the driving battery. The results also demonstrate that a small energy storage capacity, which is a disadvantage of the electric vehicle, is largely compensated without increasing the complexity of the electric vehicle.

1. Introduction

Due to emitting zero greenhouse gas, electric vehicles (EVs) are completely eco-friendly [1]. In addition, heat loss of a motor in EVs is very low as compared to that of internal combustion engines, so a high driving efficiency can be expected. Accordingly, the interest in the EVs that can solve the problems of internal combustion engine cars has been increasing worldwide. However, energy storage capacity of batteries, energy storage devices of EVs, is significantly smaller as compared to that of internal combustion engines [1,2]. Therefore, commercialization of EVs requires that EVs have efficient use and management of energy.

The main components of an EV include inverter, cluster, battery, driving motors, and battery management system (BMS) (see Figure 1). Due to the characteristics of the EVs, the most driving motor is a permanent magnet synchronous motor (PMSM) which has a high starting torque and a high efficiency [3,4]. To briefly describe each component, the battery is an energy storage device for driving EVs, while the BMS protects the battery and is a system for efficient energy use and charging [5]. The inverter converts DC voltage into AC voltage, while the PMSM generates driving force by using the converted AC voltage. When the PMSM operates as a generator, back EMF of the PMSM is converted from AC voltage to DC voltage through freewheeling diodes of the inverter [5].

*Corresponding Author: Jongsun Ko, Email: jsko@dankook.ac.kr

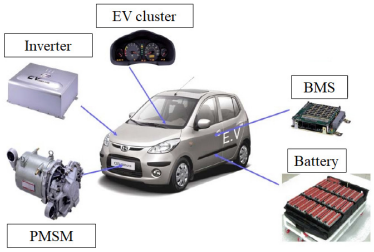


Figure 1: Main components of EVs

The PMSM may be operated as the generator when driving downhill or in coasting operation. The generated electrical energy can be used in two ways. Firstly, a method commonly called a dynamic braking consumes electric energy with resistances mounted in the EVs or eddy currents. However, a disadvantage of this method is an increase of heat inside, because electric energy dissipates into heat energy [6-8]. Secondly, a method called the regenerative energy recovery returns the generated electrical energy to the power supply system [6]. Therefore, the small energy storage capacity, which is the disadvantage of the EVs, can be compensated by using a regenerative energy recovery system (RERS) [5].

The conventional RERS recovers small energy. It increases the complexity of the EV system by using a DC/DC converter to charge the regenerative energy into supercapacitors. [5,6,9]. In

addition, a maximum energy recovery switching scheme is developed [6]. This system is charging supercapacitors using switches at the bottom of the inverter. At present, several RERS studies are underway.

In this paper, a very simple new RERS charged in a driving battery with a high capacity and a high nominal voltage through a L-C resonance circuit in EVs is proposed. As the circuit has only three capacitors and three IGBTs, the proposed system is simple. The proposed system is then analyzed by simulations using Matlab Simulink and experiments using a remodeled commercial EV.

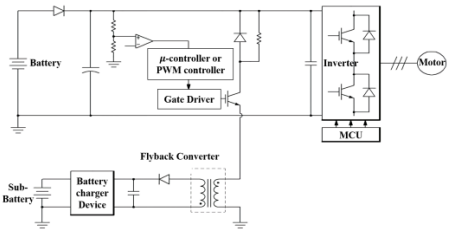


Figure 2: Regenerative energy control device for the motor driving system (Korean Patent)

2. Conventional RERS

An example of the RERS is shown in Figure 2. The battery on the upper side is used for the driving motor on the right. The flyback DC/DC converter is used to reduce the generated voltage of the motor [5]. The reduced voltage charges a sub-battery of a low capacity and a low nominal voltage. There are three disadvantages to this method. Firstly, it can be used only when the generated voltage is higher than chargeable voltage. Secondly, it cannot be used in the EV without the sub-battery. Thirdly, the regenerative power is small and has a small braking force, because it is charged in the sub-battery of a low capacity.

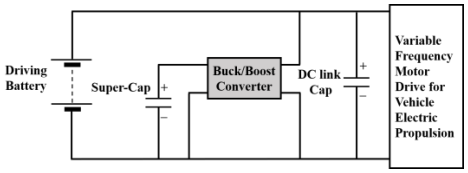


Figure 3: RERS using a supercapacitor

An example of the RERS using the supercapacitors is shown in Figure 3 [10]. When braking, the energy is charged in the supercapacitor using the buck converter and, when driving, the energy is discharged in the supercapacitor using the boost converter. Depending on the type of the RERS, the energy stored in the supercapacitor is either used directly for driving or is charged by the battery [6,10,11]. The use of DC/DC converters increases the complexity of the EV.

On the other hand, the proposed system has two advantages over the system shown in Figure 2. The first advantage is that even if the generated voltage is lower than the chargeable voltage, charging is possible by controlling the voltage through the L-C resonance control technology. Furthermore, the second advantage is that sufficient regenerative energy can be obtained because the proposed system is charging the driving battery instead of the sub-battery. Compared to the system in Figure 3, the proposed system is advantageous in the EV complexity due to its simple hardware configuration.

3. Proposed RERS Analysis

3.1. Simple Hardware Structure

Figure 4 shows a conceptual diagram of the proposed RERS. As shown in Figure 4, since the DC link voltage is short-circuited with the driving battery, the driving battery is charged only when the DC link voltage is upper than the driving battery. However, if the DC link voltage is higher than the maximum voltage of the driving battery, the BMS disconnects the driving battery to protect the battery. Therefore, the driving battery can be charged only when magnitude of the DC link voltage is within a certain range depending on the driving battery condition. In order to control DC link voltage, capacitors of C_1, C_2, C_3 are used to generate resonance in a circuit consisting of inductors and resistors of the PMSM, and a resonance rate is changed by three IGBTs above the capacitors [5]. For the purpose of three-phase equilibrium, the capacitance of C_1, C_2, C_3 is selected equally even in the resonance.

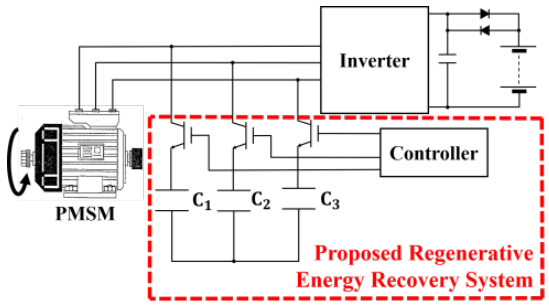


Figure 4: Conceptual diagram of the proposed system

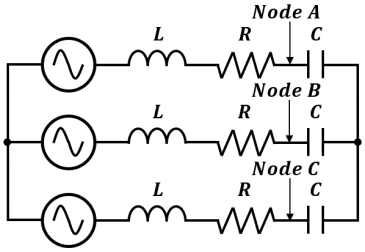


Figure 5: PMSM equivalent circuit for L-C resonance

3.2. Resonance Frequency Analysis

Figure 5 shows a PMSM equivalent circuit for L-C resonance. The L and the R are equivalent phase leakage inductance and resistance of the PMSM, and voltage sources are a three-phase voltage source generated by the PMSM. The DC link voltage varies in proportion to the voltage of nodes A, B, and C. The resonance is caused by using the capacitors in each phase for varying the voltages at nodes A, B, and C. As the circuit is a balanced three-phase circuit, the nodes A, B, and C all have same equivalent impedance. The equivalent impedance of each node is obtained as shown in Equation (1).

$$
\begin{aligned}
Z &= (j\omega L + R) \,||\, \left(\frac{1}{j\omega C} + \alpha||\alpha\right) \\
&= \frac{\mathrm{R}(C^2R^2\omega^2 + C^2L^2\omega^4 - 2CL\omega^2 + 3)}{3\{C^2R^2\omega^2 + (CL\omega^2 - 1)^2\}} \\
&+ \mathrm{j}\frac{\{CR^2(CL\omega^2 - 2) + L(CL\omega^2 - 3)(CL\omega^2 - 1)\}\omega}{3\{C^2R^2\omega^2 + (CL\omega^2 - 1)^2\}}
\end{aligned} \tag{1}
$$

with:

$$\alpha = j\omega L + R + \frac{1}{j\omega C}$$

where ω is an electrical angular frequency. The frequency at which the imaginary part of the impedance is zero and the voltage response is the largest is a resonance frequency. There are five solutions where the imaginary part becomes zero, because degree of ω is five. The solutions are shown in Equation (2).

$$\omega = 0$$
$$\omega = \pm\sqrt{\frac{-R^2C + 4L - \sqrt{R^4C^2 + 4L^2}}{2L^2C}} \quad (2)$$
$$\omega = \pm\sqrt{\frac{4L - CR^2 + \sqrt{R^4C^2 + 4L^2}}{2L^2C}}$$

The electrical angular frequency is always higher than zero, unless the EV is stopped; furthermore, assuming Equation (3),

$$R^2C \approx 0 \quad (3)$$

the solutions which cause the imaginary part to be zero are as follows (see Equation (4)):

$$\omega \approx \frac{1}{\sqrt{LC}}, \sqrt{\frac{3}{LC}} \quad (4)$$

Among the two frequencies, the frequency with a larger voltage response is the resonance frequency. The resonance frequency is selected considering the vehicle's main speed, wheel radius, gear ratio, number of poles [5]. The speed of the motor is 1km/h and the electric frequency of the motor to match the resonance frequency is : f_e=1km/1h*1h/3600sec*1000m/1km *1/2πrm*poles/2*gear ratio. The capacitance is calculated from this electric frequency(f_e) and equation (4) [5]. In order to achieve the resonance effect frequently, it is advantageous to match the main driving speed of the vehicle with the resonance frequency.

3.3. Generated Voltage Control Algorithm

A flow chart of the generated voltage control algorithm is shown in Figure 6. The flow chart shows a method of controlling the DC link voltage to the suitable voltage for the driving battery charging according to the condition of the DC link voltage. If there is no a wide gap between the DC link voltage V_{DC} and flag voltage V_{flag} , voltage control is possible by the resonance rate. Therefore, the V_{DC} is compared with the minimum controllable voltage $V_{c.min}$ and the maximum controllable voltage $V_{c.max}$. If the V_{DC} is between the $V_{c.min}$ and the $V_{c.max}$, compare with the V_{flag}. If the V_{DC} is higher, a reverse PI duty control is performed. A pulse width modulation of 1kHz is then generated through a pulse generator and used as IGBT gate input signals. There are the three IGBTs in the proposed RERS in Figure 4. In the opposite case, the PI duty control is performed.

The reverse PI duty control in the proposed system sets P gain and I gain to negative values, and the PI duty control sets P gain and I gain to positive values. Therefore, the PI duty control increases the duty, as the difference of DC link voltage at the reference voltage increases, and the reverse PI duty control decreases the duty. The adjusted duty increases or decreases the DC link voltage to match the reference voltage.

If the reference voltage of the PI duty control $V_{ref(PI)}$ and the reference voltage of the reverse PI duty control $V_{ref(RPI)}$ are equal to V_{flag}, the ripple of the V_{DC} makes comparison between V_{DC} and V_{flag} very difficult. Therefore, the $V_{ref(RPI)}$ is set above the V_{flag} and below the battery maximum voltage and the $V_{ref(PI)}$ is set below the V_{flag}.

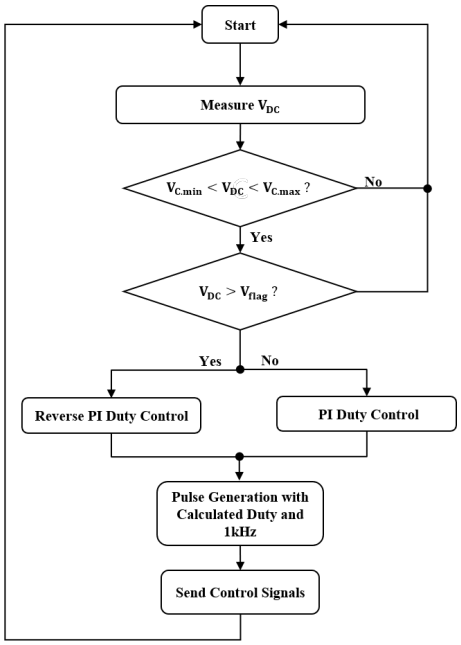


Figure 6: Flow chart

4. Simulations

Three simulations have been performed. First, a simulation of frequency response has been performed. Characteristic of frequency response is that the equivalent circuit of a PMSM's stator is taken into account. The second simulation has been performed to verify the generated voltage control algorithm. The characteristic to verify the generated voltage control algorithm is that the DC link voltage response is fast because of low inertia, low viscous friction and short battery response time. Using this fast response, the proposed system algorithm is verified. The third simulation has been performed taking into account the remodeled commercial EV. The characteristics considered the EV are that the DC link voltage response is slow, because the EV has high inertia, high viscous friction, and long battery response time.

4.1. Frequency Response of Equivalent Circuit

Table 1: 3kW PMSM Parameters Used for the Remodeled Commercial EV

Parameter	Value	Unit
Rated Power	3	kW
Rated Current	17.2	A_{rms}
Rated Speed	3000	rpm
Rated Torque	9.54	Nm
Stator Phase Resistance	0.102	Ω
Stator Phase Inductance	1.22	mH
Poles	8	-
Inertia	0.28	kgm^2
Viscous friction	0.026	Nms

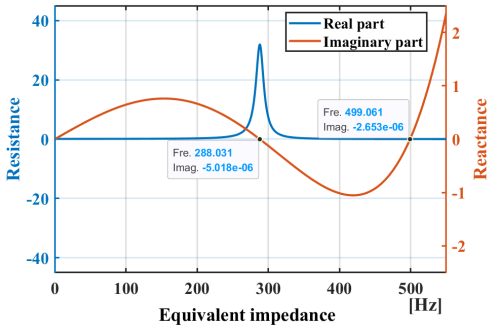


Figure 7: Frequency response of PMSM equivalent circuit for L-C resonance

The simulator for the frequency response with the capacitors and without the capacitors is shown in Figure 8. If capacitors in a red box in Figure 8 is removed, the simulator is a common 3-phase rectifier circuit. The opposite case is the L-C resonance circuit. This simulator can find the resonance frequency, which is either 288Hz or 499Hz, and also shows the resonance effect.

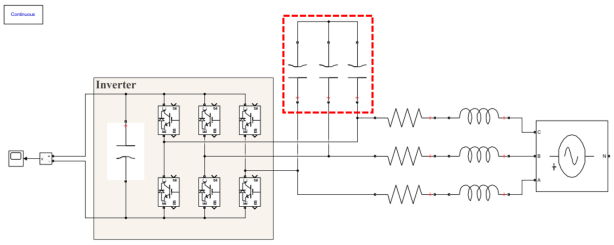


Figure 8: Simulator for frequency response

In this simulation, the DC link voltage for frequency is shown in Figure 9. The frequency of the highest voltage response is 288Hz in the L-C resonance circuit. Therefore, the resonance frequency is 288Hz. From 1Hz to 416Hz, the DC link voltage of the circuit with the capacitors is higher than the DC link voltage of the circuit without capacitors. In addition, after 416Hz, the DC link voltage of the circuit with the capacitors is lower than the DC link voltage of the circuit without the capacitors. In PMSM, generated voltage is generally proportional to frequency. When the frequency and generated voltage are low, the generated voltage can be raised using the L-C resonance circuit [5]. In contrast, when the frequency is high, and the generated voltage is high, the generated voltage can be reduced using the L-C resonance circuit. In this way, it is possible to control voltage, since the resonance rate is changed when the capacitors and the IGBTs switching duty are adjusted.

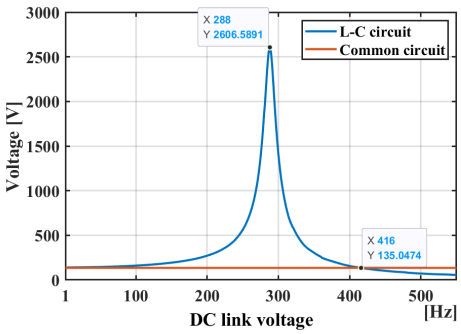


Figure 9: Simulation results of frequency response

4.2. Generated Voltage Control Algorithm Verification

The simulator to verify the generated voltage control algorithm is shown in Figure 10. The parameters used in this simulator are shown Table 2. Due to the high back EMF of the PMSM, RCD snubber is used to reduce the switching stress of the IGBTs. Due to the small inductance of 4.5kW PMSM, voltage rise using resonance is low at the desired speed. To maximize the voltage rise, the inductance (15mH) is added in front of the capacitance (1.25mF).

Table 2: Parameters Used in the Algorithm Verification

Item	Parameter	Value	Unit
PMSM	Rated Power	4.5	kW
	Rated Current	17.9	A_{rms}
	Rated Speed	2000	rpm
	Rated Torque	21.5	Nm
	Stator Phase Resistance	0.2	Ω
	Stator Phase Inductance	2.625	mH
	Poles	6	-
	Inertia	0.00369	kgm^2
	Viscous friction	10	μNms
Driving Battery	Initial SoC	84	%
	Rated Power	3	kW
	Rated Voltage	295	V
	Maximum Voltage	340	V
	Minimum Voltage	224	V
	Response Time	30	sec
Proposed System	Phase Capacitance	1.25	mF
	Phase Inductance	15	mH
	Snubber Resistance	500	Ω
	Snubber Capacitance	4.7	nF
	V_{flag}	332	V
	$V_{c.min}$	280	V
	$V_{c.max}$	400	V
	$V_{ref(RPI)}$	335	V
	$V_{ref(PI)}$	330	V

Simulation results without the algorithm are shown in Figure 11, while simulation results with the algorithm are shown in Figure 12. The torque input from 0 to 15 seconds is -5Nm, and torque input from 15 to 30 seconds is -10Nm. A flag in Figures 11 and 12 is a variable for distinguishing the control (1 : reverse PI duty control, 2 : PI duty control, 3 : no control). In a voltage waveform

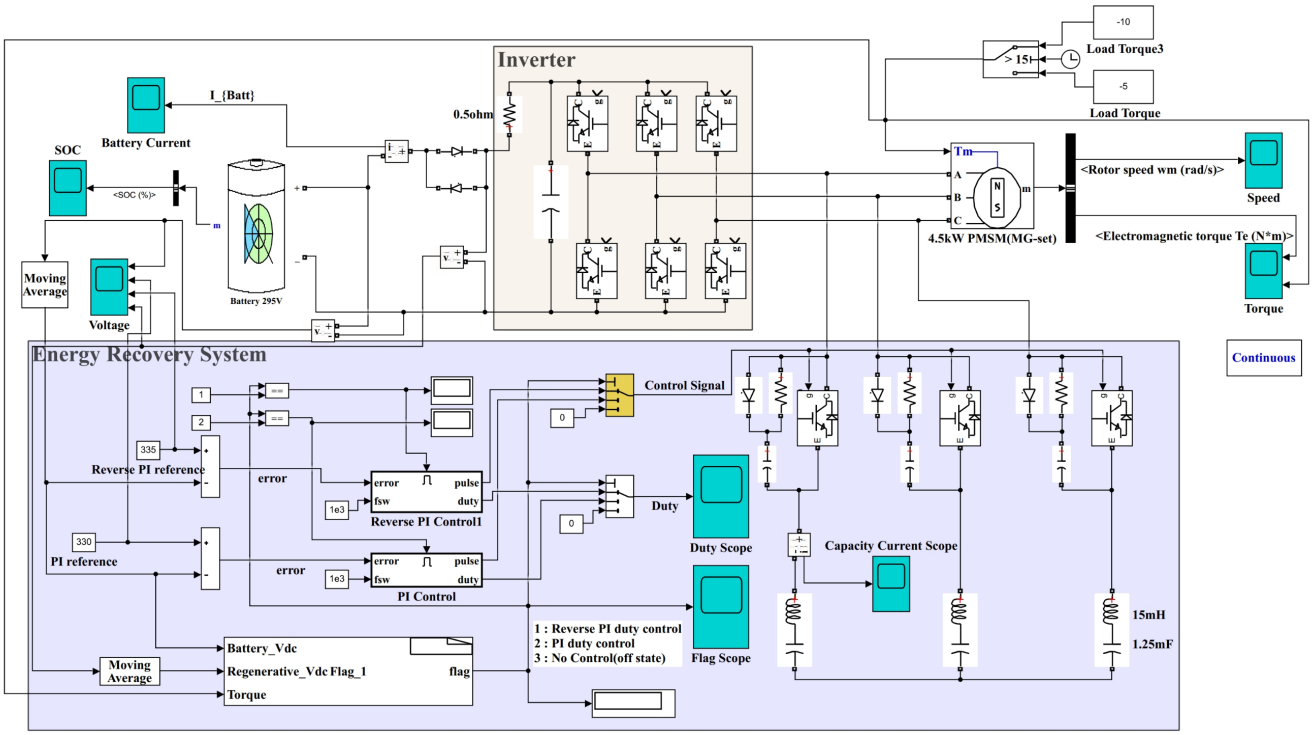


Figure 10: Simulator for the generated voltage control algorithm verification

of Figure 12 from 0 to 15 seconds, the PI duty control (flag 2) is performed, because the V_{DC} is lower than the V_{flag}. In order to raise the generated voltage, the duty ratio of the PI duty controller steady state output is 1. The voltage and current waveforms in Figure 12 are higher than those in Figure 11. In the voltage waveform of Figure 12 from 15 to 30 seconds, the reverse PI duty control (flag 2) is performed, because V_{DC} is higher than V_{flag}. The resonance rate is changed by the reverse PI duty control so that V_{DC} is controlled to $V_{ref(RPI)}$ in voltage in Figure 12. On the other hand, in the voltage in Figure 11, the battery voltage V_{batt} rises above the battery maximum voltage. In the EV, when the battery voltage rises above the battery maximum voltage, the BMS disconnects the DC link capacitor from the battery. Therefore, battery charge current I_{BC} does not flow when EVs are applied. Because V_{batt} in Figure 12 does not exceed the battery maximum voltage, the battery charge current I_{BC} flows when the proposed system is applied to the EVs.

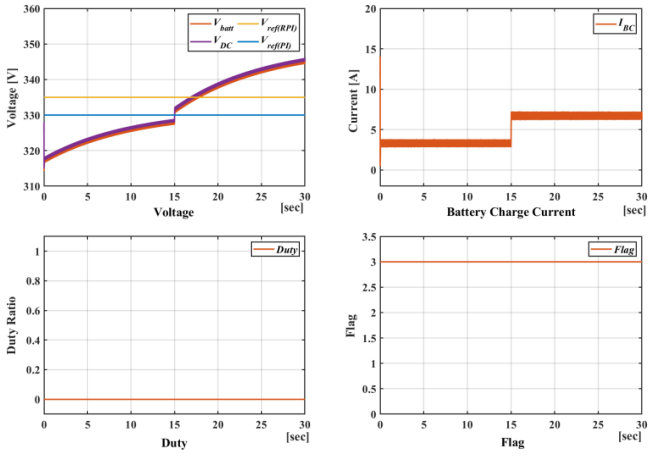


Figure 11: Simulation results without the voltage control algorithm on torque variation

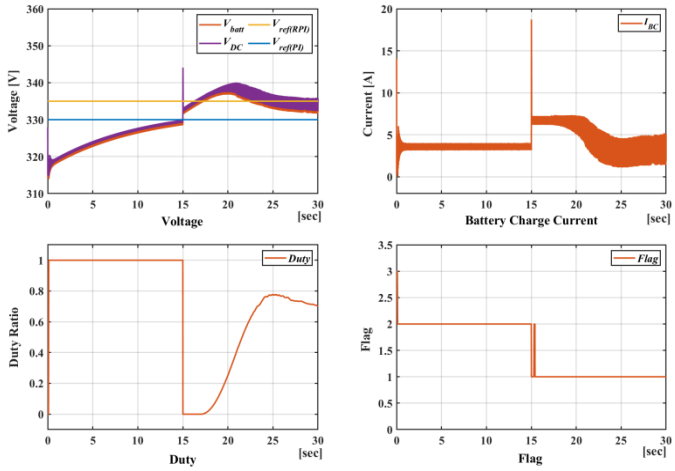


Figure 12: Simulation results without the voltage control algorithm on torque variation

4.3. Energy Recovery Rate Considered the Remodeled EV

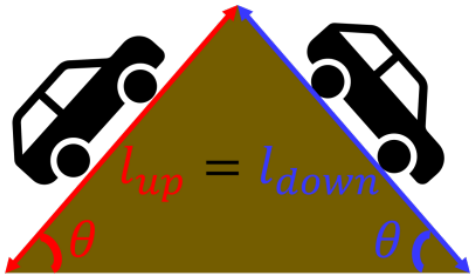


Figure 13: Road environment measured by the energy recovery rate

An energy recovery rate η is defined by Equation (5). E_{BC} is battery charge energy, and E_{BD} is battery discharge energy.

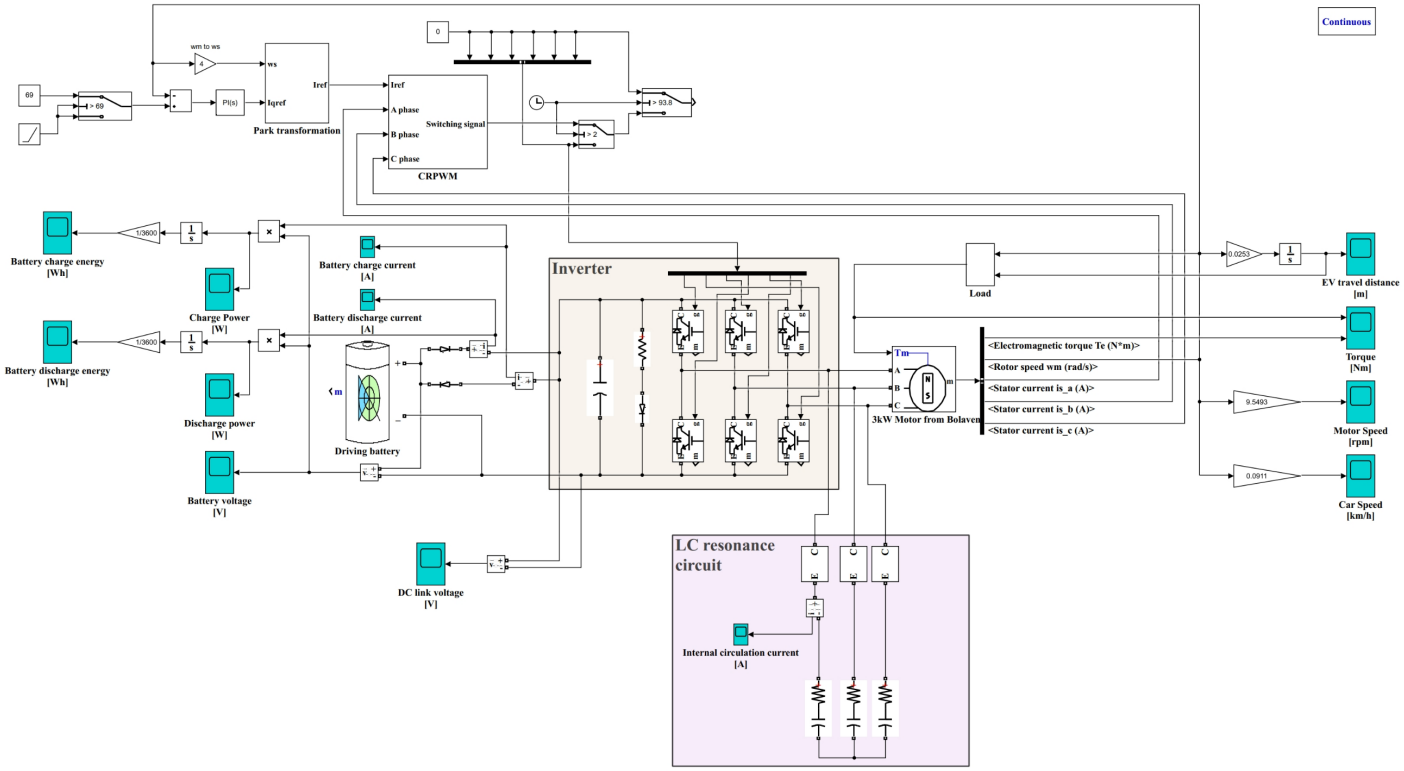


Figure 14: Simulator considered the remodeled commercial EV

$$\eta = \frac{E_{BC}}{E_{BD}} \times 100 \tag{5}$$

Measurement road environment is a road with the same travel distance and the same slope as shown in Figure 13. Since the EV input by the slope is the force of $\mathrm{mgsin}(\theta)$, this torque is put into the mechanical input of the PMSM in the simulation.

Travel distance of the EV D_{EV} is calculated using Equation (6). When the EV moved 150m, PMSM is stopped so that the same travel distance is obtained.

$$D_{EV} = \int (\omega_m \times \frac{1}{2\pi G} \times 2\pi r) dt \tag{6}$$

A ω_m is the speed of the PMSM (rad/s), the G is the gear ratio and r is the wheel radius of the EV.

Table 3: Parameters of the Remodeled Commercial EV

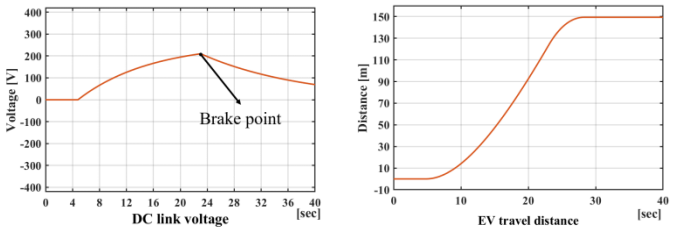
Item	Parameter	Value	Unit
Driving Battery	Initial SoC	10.5	%
	Rated Power	3	kW
	Rated Voltage	295	V
	Maximum Voltage	340	V
	Minimum Voltage	224	V
	Response Time	500	sec
Proposed System	Phase Capacitance	250	μF
Power Transmission Device	Gear	9.3	–
Road Environment	Torque	13.5	Nm
	Distance	150	m
	Wheel Radius	0.235	m

A simulator considered the remodeled commercial EV is shown in Figure 14. Applying the environment in Figure 13, the parameters in Tables 1 and 3 to this simulator, the 3kW PMSM cannot run the high speed needed for the reverse PI duty control. For that reason, only the PI duty control has been performed. Also, due to the slow DC link voltage response, the duty of the PI duty control output is 100%. Therefore, the three IGBTs of the proposed system are replaced by shorting a emitter and a collector. The three IGBTs are short-circuited only in the downhill condition where the PMSM can act as a generator. The resistance of the L-C resonance circuit is equivalent to lead resistance, which is very small. The PMSM is driven by speed control in the uphill condition.

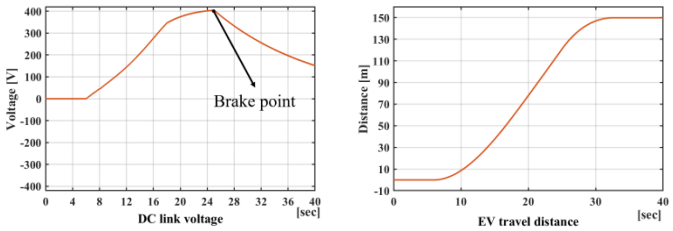
Two simulations considered the remodeled commercial EV have been performed. Firstly, to see change in the generated voltage, existence and non-existence of the proposed system have been simulated after removing the driving battery. Secondly, to see change in the energy recovery rate, existence and nonexistence of the proposed system have been simulated after installing the battery.

The generated DC link voltage and the remodeled EV travel distance after the proposed system is removed are shown in Figure 15 (a). Before connecting the battery, the simulation results in the capacitor model without the battery to see the controllability with the voltage that can be charged to the battery through the proposed system. The generated DC link voltage and the remodeled EV travel distance after the proposed system is installed are shown in Figure 15 (b). The remodeled EV travels the same distance of 150m, but the DC link voltage in Figure 15 (b) nearly doubles as compared to the DC link voltage in Figure 15 (a). The DC link voltage in Figure 15 (a) is always lower than the minimum voltage of the driving battery, so the battery cannot be charged. However,

the DC link voltage in Figure 15 (b) can be charged, because there is a section higher than the minimum voltage of the driving battery. At brake points in Figure 15 (a) and Figure 15 (b), a mechanical brake is activated, after which the DC link voltage is reduced.



(a) Without the proposed system



(b) With the proposed system

Figure 15: Downhill simulation results without the driving Battery

Battery voltage, discharge current, discharge energy, and the remodeled EV travel distance in the driving uphill condition are shown in Figure 16. In this condition, the driving battery is discharged. 32.8689Wh is consumed while traveling 150m. As the driving battery's energy is consumed, the battery voltage in Figure 16 decreases. The battery discharge current in Figure 16 is displayed because the remodeled EV is initially driven at the accelerated speed, while the remodeled EV is driven at the constant speed after 18 seconds.

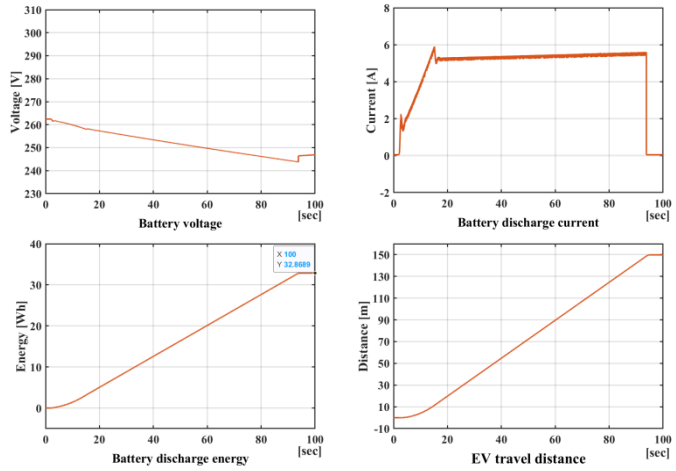
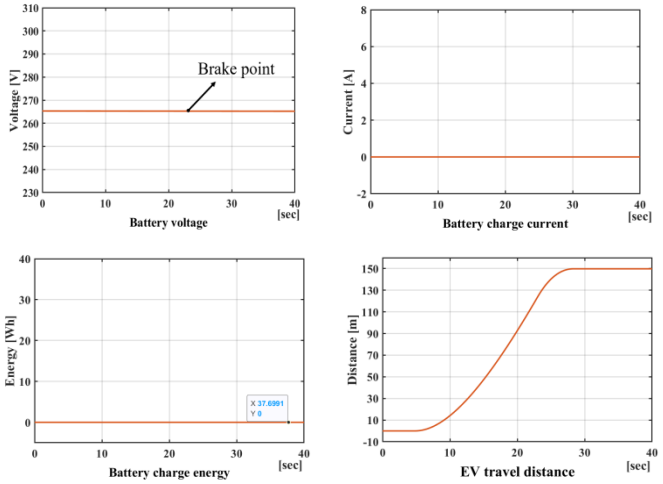


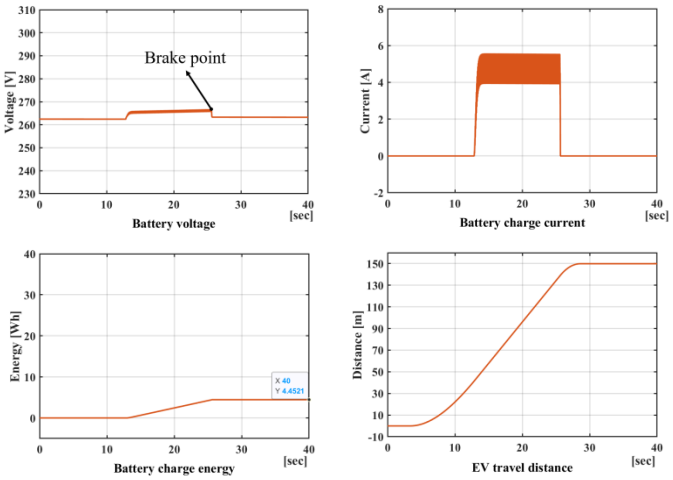
Figure 16: Uphill simulation results with the driving battery

Battery voltage, charge current, charge energy, and the remodeled EV travel distance in the driving downhill condition without the proposed system are shown in Figure 17 (a). Although the remodeled EV traveled 150m, the energy is not recovered, as the generated voltage is lower than the battery voltage.

Battery voltage, charge current, charge energy and remodeled EV travel distance in the driving downhill condition with the proposed system are shown in Figure 17 (b). While the remodeled EV traveled 150m, the energy is recovered because the generated voltage is lower than the battery voltage. Therefore, the battery voltage and the charge current raise. At the brake point, the mechanical brake is activated, after which the battery stops charging. The charged energy in the battery is 4.451Wh. The energy recovery rate is computed to 13.54% using Equation (5).



(a) Without the proposed system



(b) With the proposed system

Figure 17: Downhill simulation results with the driving battery

Battery voltage, charge current, charge energy, and the remodeled EV travel distance in the driving downhill condition without the proposed system are shown in Figure 17 (a). Although the remodeled EV traveled 150m, the energy is not recovered, as the generated voltage is lower than the battery voltage.

Battery voltage, charge current, charge energy and remodeled EV travel distance in the driving downhill condition with the proposed system are shown in Figure 17 (b). While the remodeled EV traveled 150m, the energy is recovered because the generated voltage is lower than the battery voltage. Therefore, the battery voltage and the charge current raise. At the brake point, the mechanical brake is activated, after which the battery stops charging. The charged energy in the battery is 4.451Wh. The energy recovery rate is computed to 13.54% using Equation (5).

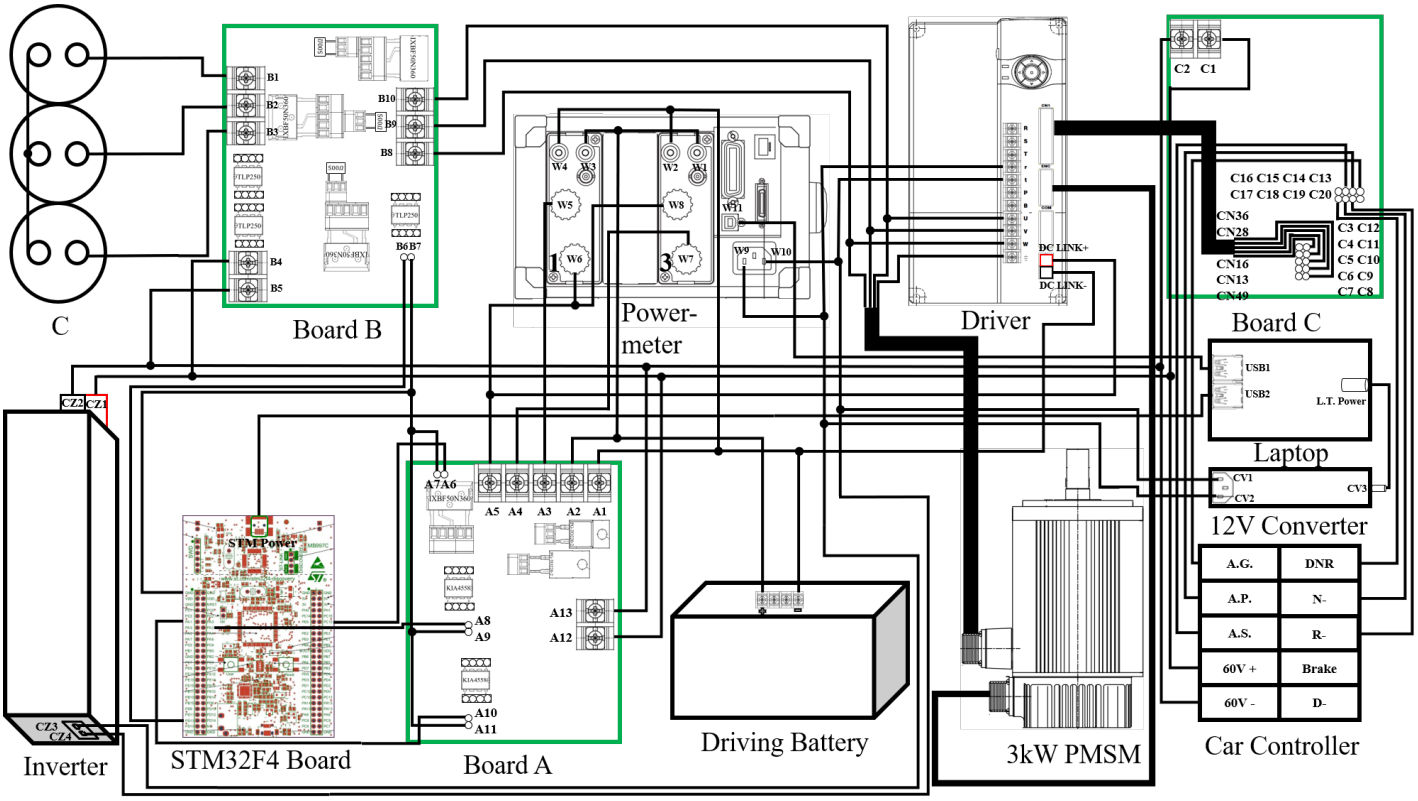


Figure 19: Connection diagram of the remodeled commercial EV

5. Experiment

A small commercial EV has been replaced with the 3kW PMSM, motor driver, gear and driving battery; then the proposed system has been installed in a small commercial EV. Interior and exterior of the remodeled commercial EV are shown in Figure 18. The upper left is a front part of the remodeled EV, while the upper right is a picture of installing the replaced 3kW PMSM. The bottom left shows equipment installed in a back seat of the remodeled EV. This equipment includes a board C for connecting the remodeled EV, motor driver, board A and B for the proposed system, inverter for driving the motor driver, laptop, power meter, and so on. In the bottom right, there are the power meter and laptop for measuring the energy recovery rate, and three capacitors of the proposed system.



Figure 18: The remodeled commercial EV

A wiring diagram of the equipment has mounted on the remodeled EV is shown in Figure 19. A car controller, which send signals from the remodeled EV, has been connected to the motor driver via the board C.

A map and road at the experimental site place in Figure 13 are shown in Figure 20. Two experiments have been performed. First, to see the change in the generated voltage, existence and nonexistence of the proposed system have been experimented with after removing the driving battery. Second, to see the change in the battery charge energy, existence and nonexistence of the proposed system have been experimented with after installing the driving battery. The parameters used in the experiments are shown in Tables 1 and 3.

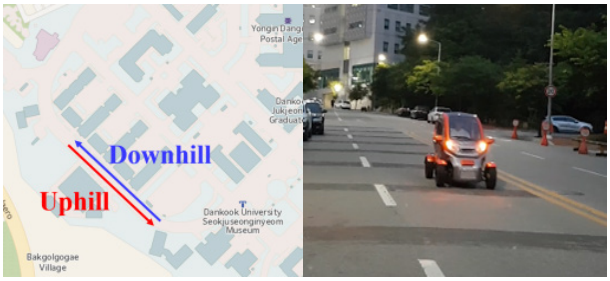
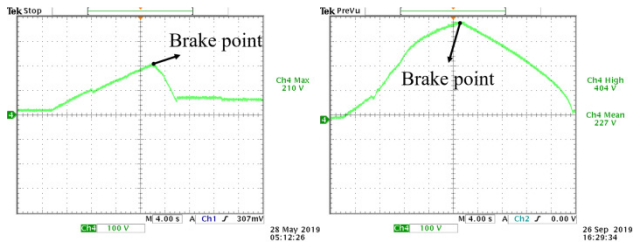


Figure 20: Map and road of the experiment place

In downhill experiments, the low slope of the road cannot output the high speed needed for the reverse PI duty control. For that reason, only the PI duty control has been performed in the experiments. Also, due to the slow DC link voltage response, the duty of the PI duty control output is 100%. Therefore, the three IGBTs of the proposed system have been replaced by shorting a emitter and a collector.

5.1. Energy Recovery Rate Considered the Remodeled EV

Figure 21 shows experimental results without the driving battery. Figure 21 (a) shows DC link voltage without the proposed system and (b) shows DC link voltage with the proposed system. Comparing the (a) and (b) in Figure 21, the DC link voltage almost doubled. The driving battery cannot be charged in Figure 21 (a), because the DC link voltage is always lower than the minimum voltage of the driving battery. However, the DC link voltage in Figure 21 (b) can be charged, because there is a section higher than the minimum voltage of the driving battery. At brake points in Figure 21 (a) and (b), the mechanical brake is activated, after which the DC link voltage is reduced.



(a) Without the proposed system (b) With the proposed system

Figure 21: DC link voltage of downhill experiment without the driving battery

Battery voltage, discharge current, discharge power, and discharge energy in the driving uphill condition are shown in Figure 22. As much as the driver activates the accelerator, the battery discharge current and discharge power is shown. 32.878Wh is consumed while traveling 150m. As the driving battery's energy is consumed, the battery voltage is reduced.

The results on the driving downhill condition without the proposed system are shown in Figure 23 (a). Since the generated DC link voltage is lower than the battery voltage, the energy is hardly recovered.

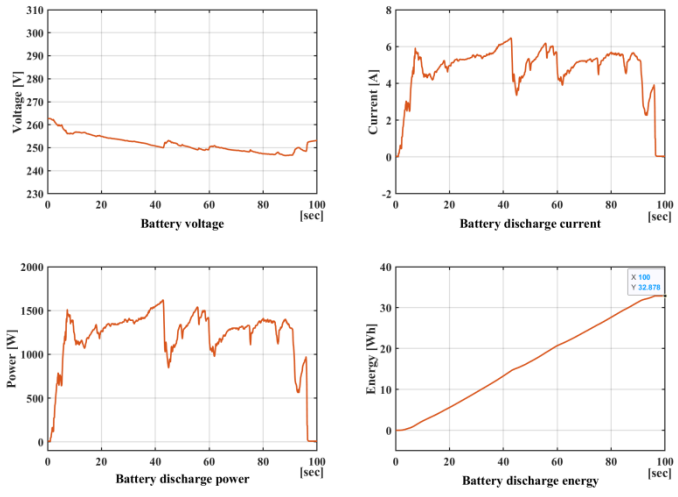
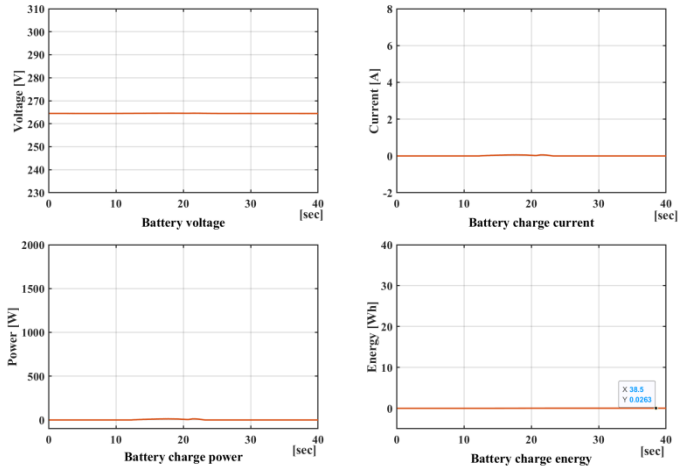
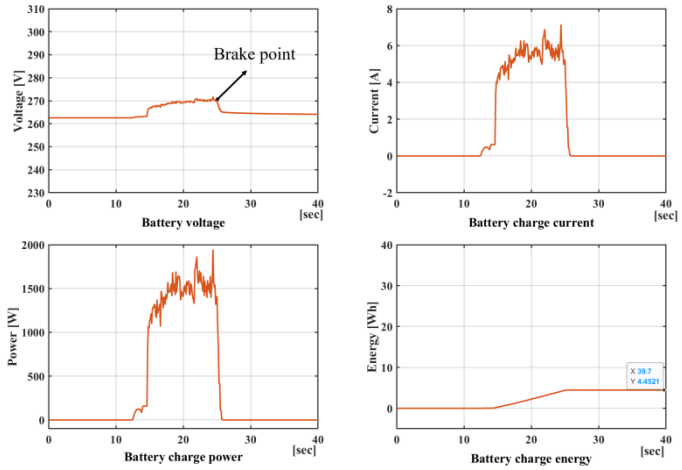


Figure 22: Uphill experiment results with the driving battery

Experimental results for the driving downhill condition with the proposed system are shown in Figure 23 (b). The generated DC link voltage is higher than the battery voltage. Therefore, the energy is charged to the battery, increasing the battery voltage and charge current and showing the charge power in Figure 23 (b). At the brake point in Figure 23 (b), the mechanical brake is activated, after which the generated voltage reduces, and the battery charge stops. The charge energy in the battery is 4.4521Wh. The energy recovery rate is computed to 13.54% using Equation (5). This result complements the small energy storage capacity of the EVs, increasing travel distance on a single charge.



(a) Without the proposed system



(b) With the proposed system

Figure 23: Downhill experiment results with the driving battery

6. Conclusion

In this paper, a very simple new RERS was proposed to complement the EVs with a low energy storage capacity. The conventional RERS uses DC/DC converters to increase the complexity of the EV or has a small recovery area. The proposed system does not increase the complexity of the EVs, as the hardware is very simple. Furthermore, the recovered energy is large, because the generated voltage is controlled to the optimal level for the driving battery.

Three major analyses were performed. Firstly, the proposed system was analyzed. The results of this analysis showed that the circuit is very simple, and resonance frequency was formulated. Secondly, the simulation of the proposed system was performed, which allowed us to test the formulated resonance frequency and the generated voltage control algorithm. In addition, the energy recovery rate amounted to 13.54% in the simulation considering

the remodeled commercial EV. Third, the using the remodeled commercial EV was performed. In the experiment, the energy recovery rate amounted to 13.54%. The results showed that the proposed system can very simply significantly compensate for the small energy storage capacity, which is a disadvantage of the EVs.

Acknowledgments

The present research was supported by the research fund of Dankook University in 2020.

References

[1] C. C. Chan, "The State of the Art of Electric, Hybrid, and Fuel Cell Vehicles," in Proceedings of the IEEE, vol. 95, no. 4, pp. 704-718, Apr. 2007. https://doi.org/10.1109/JPROC.2007.892489

[2] C. Zheng, W. Li and Q. Liang, "An Energy Management Strategy of Hybrid Energy Storage Systems for Electric Vehicle Applications," in IEEE Transactions on Sustainable Energy, vol. 9, no. 4, pp. 1880-1888, Oct. 2018. https://doi.org/10.1109/TSTE.2018.2818259

[3] G. Pellegrino, A. Vagati, B. Boazzo and P. Guglielmi, "Comparison of Induction and PM Synchronous Motor Drives for EV Application Including Design Examples," in IEEE Transactions on Industry Applications, vol. 48, no. 6, pp. 2322-2332, Nov. 2012. https://doi.org/10.1109/TIA.2012.2227092

[4] J. Nerg, M. Rilla, V. Ruuskanen, J. Pyrhönen and S. Ruotsalainen, "Direct-Driven Interior Magnet Permanent-Magnet Synchronous Motors for a Full Electric Sports Car," in IEEE Transactions on Industrial Electronics, vol. 61, no. 8, pp. 4286-4294, Aug. 2014. DOI:10.1109/TIE.2013.2248340

[5] K. Park, S. Jung, S. Kim and J. Ko, "Generated Voltage Control of EV PMSM for Maximizing Energy Recovery Rate using Switched Resonant L-C," 2019 14th Annual Conference System of Systems Engineering (SoSE), Anchorage, AK, USA, May 2019, pp. 195-199. https://doi.org/10.1109/SYSOSE.2019.8753832

[6] R. E. Hellmund, "Regenerative Braking of Electric Vehicles," in Transactions of the American Institute of Electrical Engineers, vol. XXXVI, pp. 1-78, Jan. 1917. https://doi.org/10.1109/T-AIEE.1917.4765458

[7] Sung-Chul Jung, In-Sik Yoon, and Jong-Sun Ko, "Electromagnetic Retarder´s Power Recovery Device and Voltage Control," in The Transactions of Korean Institute of Power Electronics, vol. 21, no.5, pp. 396-403, Oct. 2016.

[8] Sung-Chul Jung, Jong-Sun Ko, "A Study on Electromagnetic Retarders Power Recovery System and Regenerating Voltage Control," in The Transactions of The Korean Institute of Electrical Engineers, vol. 66, no. 8, pp. 1207-1214, Aug. 2017.

[9] P. J. Grbovic, P. Delarue, P. Le Moigne and P. Bartholomeus, "The Ultracapacitor-Based Regenerative Controlled Electric Drives With Power-Smoothing Capability," in IEEE Transactions on Industrial Electronics, vol. 59, no. 12, pp. 4511-4522, Dec. 2012. https://doi.org/10.1109/TIE.2011.2181129

[10] S. Lu, K. A. Corzine and M. Ferdowsi, "A New Battery/Ultracapacitor Energy Storage System Design and Its Motor Drive Integration for Hybrid Electric Vehicles," in IEEE Transactions on Vehicular Technology, vol. 56, no. 4, pp. 1516-1523, July 2007. https://doi.org/10.1109/TVT.2007.896971

[11] J. Cao and A. Emadi, "A New Battery/UltraCapacitor Hybrid Energy Storage System for Electric, Hybrid, and Plug-In Hybrid Electric Vehicles," in IEEE Transactions on Power Electronics, vol. 27, no. 1, pp. 122-132, Jan. 2012. https://doi.org/10.1109/TPEL.2011.2151206

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 67-90 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Multi Biometric Thermal Face Recognition Using FWT and LDA Feature Extraction Methods with RBM DBN and FFNN Classifier Algorithms

Evan Hurwitz, Chigozie Orji*

The University of Johannesburg, Department of Electrical and Electronic Engineering Science, Johannesburg, South Africa

ARTICLE INFO

Article history:
Received: 12 September, 2019
Accepted: 14 November, 2019
Online: 22 November, 2019

Keywords:
Thermal imaging
Soft-biometrics
an ensemble of classifiers
Fast Wavelet Transform
Linear Discriminant Analysis
Restricted Boltzmann Machines
Deep Belief Networks
Feed Forward Neural Networks
Sensing
Imaging
Security

ABSTRACT

Person recognition using thermal imaging, multi-biometric traits, with groups of feature filters and classifiers, is the subject of this paper. These were used to tackle the problems of biometric systems, such as a change in illumination and spoof attacks. Using a combination of, hard and soft-biometric, attributes in thermal facial images. The hard-biometric trait, of the shape of a head, was combined with soft-biometric traits such as the face wearing glasses, face wearing a cap/headgear, face with facial hairs, plain face, female face, and male face. These were experimented with, using images from Carl's database and Terravic Facial Infrared Database, and used to train clusters of neural network algorithms for each biometric trait. These comprised Restricted Boltzmann Machines (RBM), Deep Belief Networks (DBN), and Feed Forward Neural Networks (FFNN). After feature extraction, using Fast Wavelet Transform (FWT), and Linear Discriminant Analysis (LDA). A classification error of 0.02887, 0.038695, 0.02381, 0.024629, 0.0268, 0.02369 and 0.03 was achieved for each biometric trait, respectively. Showing that they had each been learned, and could be used through a fusion method, to improve recognition. This was demonstrated using a test image, as the user, having four of the character traits (countenance, glasses, facial hair, and gender). Then attempting to recognize each trait, one after the other, using a cross-verification method. The algorithm was seen to return test values, close to those received during the training test, for each biometric trait.

1. Introduction

Thermal imaging is a preferable choice for facial biometrics because inconsistencies due to change in illumination are not present. The human body is luminous at frequencies, such as microwave, millimeter-wave and infrared. This makes it possible for thermal sensors to passively capture images of a consistent nature at these frequencies. Which are useful for biometric system applications such as person recognition [1]. Other frequencies, which are beyond Ultraviolet, are not suitable for such applications due to health-related issues such as genetic mutation and skin cancer, arising from their prolonged exposure [2]. Figure 1 shows these frequency bands, highlighting the millimeter-wave band and that of visible light [3].

1.1. Person recognition

Person recognition has been an ever-growing work area, in visual sensor based biometrics. A field within which scientists and engineers are continually improving existing algorithms and developing new ones. These are useful in systems which detect, recognize and differentiate one type of person from the other. The following applications use biometrics with computer vision:

- Personal entertainment and assistance - Comprising video games, virtual reality, human-computer interaction, avatars, and robotics applications.
- Public enrolment and security - Some of which are driver's license, passport, voting, registration and enlistment programs. Along with welfare, identity document, passport, and voter forgery identification systems.
- Device access and authentication - Consisting of TV child proof locks, desktop access, personal device sign-in (mobile device and tablets), network and database authentication. Also, internet access, online record and portal access (medical and law enforcement) with removable storage device access.

* Chigozie Orji & Email: chris.sardius@gmail.com

- Remote monitoring and surveillance - Such as advanced video surveillance and CCTV Control in airports (drug trafficking), malls and departmental stores (shoplifting). Also conference centers, cinemas, private premises, office buildings, and convenient stores.

This paper is an extension of the research work done by Evan, Ali and Chigozie [4]. It focuses on the appraisal of biometric system security and the proposal of a protection method, using thermal imaging and multi-biometric [5] traits to harden a biometric system through fusion [6] . Galbally, Marcel and Fierrez [7] discussed ways of fooling a biometric system into, passing an improper user as a legitimate one. By presenting a counterfeit crafted version of the genuine biometric trait to the sensor. Such as, 2D face spoofing using fake fingerprints, irises, photos, videos. Also 3D face spoofing using masks made of paper, plastic or silicone. Such spoofing methods, have been used to, compromise secure systems and reduce the confidence people have in biometrics. According to Hong, Jain and Pankanti [8] the integration of additional pieces of evidence, makes a biometric system more reliable and, increases verification accuracy. Various approaches are possible, such as the integration of semblance and fingerprint, semblance and iris, and appearance and voice. For this research, additional evidence was created, using thermal images of *soft biometric* attributes, along with the *hard biometrics* of the human face. A technique which doesn't require, the use of, multiple sensor input devices to increase robustness.

The weaknesses that can exist in a recognition system arise from a hacker's awareness of the biometric information required for user authentication and his ability to falsify them. The use of hidden, authentication, evidence improves a system's robustness and hardens it against such attacks. All possible combinations, of biometric attributes, have not been exhausted. That is why the work in this research is focused on, multi-biometric person recognition, using soft and *hard biometric* traits in thermal facial images. Using a matching score and fusion approach, to combine biometric evidence and altogether boost security.

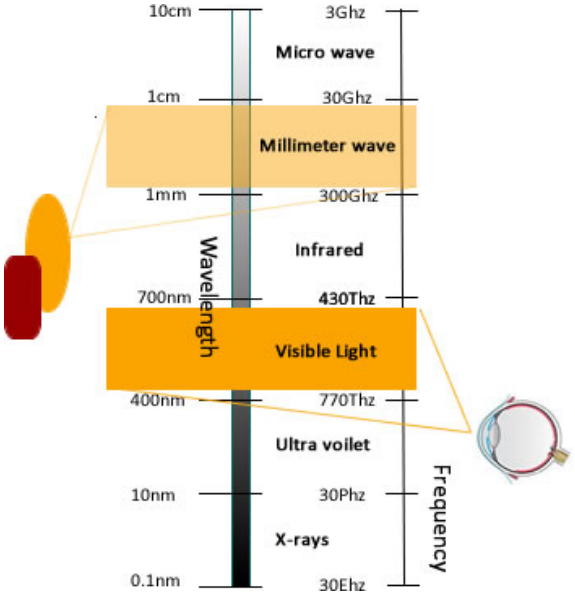


Figure 1: Imaging beyond the visible spectrum, showing the optical eye, using visible light frequencies, and thermal sensor device, using millimeter-wave frequencies

1.2. Data sets

The availability of datasets helped to further motivate the research initiative. These comprised images of different types of people captured under diverse conditions such as poses, lighting, dressing, and makeup. From these samples, for training and testing the anti-spoof algorithm to achieve reduced vulnerability, were derived. The Object Tracking and Classification Beyond the Visible Spectrum (OTCBVS) benchmark dataset collection [9], and Carl's databases [10], have provided a dataset of images, captured under diverse conditions, using thermal imaging. Which can help research work in the area of person recognition. The facial category, within the OTCBVS benchmark dataset collection: Terravic Facial IR Database, provides thermal facial images with good variability. These comprise real-life shots having diverse behavioral characteristics. While Carl's database consists of real-life facial images, captured simultaneously, using visible and invisible frequencies. They have gender variation and good behavioral attributes. Both datasets, comprise ideal thermal facial human categories and, are suitable for measuring the working of a multi-biometric person recognition system.

2. Literature Review

Biometrics is the science of knowing a person's identity, using thier physical or behavioral qualities, such as the semblance, fingerprints, voice, and iris [11]. The idea, that spurred biometrics as a science, was introduced within the field of law enforcement. When Alphonse Bertillon, police head of the criminal identification department in Paris, introduced a new way to identify criminals. Using using body measurements such as eye color, skin, hair, beard, fingers, shape, and size of the head, [12].

A classical biometric system can verify identity by:

- Capturing biometric information such as images, fingerprint and voice samples from people.
- Extract features from the information.
- Match the feature set against feature patterns in the database.
- Conclude an identity verification.

Aside from verification, information collected by a biometric algorithm can be used to ascertain additional peculiarities of the persons concerned. Traits such as hair, gender, gait, age, clothing, ethnicity, and body weight can be captured [13]. This additional data can help in bettering recognition verity in a biometric algorithm [14].

Biometrics also saw application in the field of legal work such as criminal identification and illegal migrant monitoring. It is also useful for security clearances among employees in sensitive occupations, forensics, the identification of prisoners among suspects and ex-convicts. In today's world the private sector, comprising of civilians, is beginning to add biometrics to alphanumeric methods, for person recognition [15].

2.1. Biometrics systems

A biometric system can be used for recognizing persons using the recognition of patterns in a feature vector. These can be derived from a particular corporal or behavioral attribute that the individual

possesses. Depending on the type of application a biometric system works either in *verification* or *identification mode* [16].

2.2. Biometric system errors

Multiple samples of simmilar biometric attribute, retrieved from one individual, are not always alike. Such as two exposures of a person's face, two prints of a person's left index finger and two voice samples from the same person. Effects such as changes in the user's physiology, caused by wounds, cuts, unusual dressing, moods, ambient lighting, weather, and age, can cause errors in a biometric system. The user's way of interacting with the sensor can have the same effect. As a result a matching score, in numerical terms, is used to describe the similtude between the input (captured biometric trait) and database template. Such that, the higher the score the more convinced the system is that, the two biometric data come from the same user. The system's decision is regulated by a threshold "*s",* for the matching score, such that the higher the "*s*" the more secure the system. A biometric verification system can make, two kinds of errors, a *false accept* (false match) or a false reject (false nonmatch) [17]

2.3. Hard and Soft biometrics

A person's image, normally, comprises both hard and *soft-biometric* characteristics [18]. Some hard-biometric characteristics are countenance, the shape of the head and gait. While *soft-biometric* characteristics are features such as glasses, hair, hat, clothes, gender and other occlusions that an individual may have on.

2.3.1. Hard-biometrics

Memory-based authentication methods, such as usernames and passwords, can easily be compromised if written down in plain text and such text files are copied. Such authentication systems can be hardened by capturing and verifying a *hard biometric* trait along with the alphanumeric information. *Hard biometric* methods are more dependable than alphanumeric authentication schemes. They also have several advantages that alphanumeric methods do not grant. *Hard biometric* authentication involves the establishment of a person's identity with, the use of physical characteristics. These are traits such as the face, iris, fingerprint, hand geometry, gait, voice, etc. Unlike alphanumeric tokens, *Hard biometric* traits cannot be copied, shared or distributed, they can also neither be lost or forgotten. *Hard biometrics* always requires that the person, being validated, be present at the time and place where access is being requested. *Hard biometric* traits are arduous to forge and a lot of money, time, access and experience, would be needed to forge them. For application-specific use each hard biometric trait has its advantages and disadvantages. The type of use, however, derives from the application requiring it. ATMs would use facial biometric traits, video surveillance: gait, while telebanking and telemarketing would use voice.

2.3.2. Soft-biometrics

Semantics, in the context of person recognition, are characteristics which people often use in describing those they meet but don't know personally. Such as stature, gender, skin color, hair color, body weight, age, race, facial features, mode of dressing and mannerisms [19]. Using such semantic data, with features from a biometric dataset, to further enhance the working of a person recognition algorithm, is called "*soft-biometrics*". It bears with it all the advantages of the hard-biometric authentication schemes, over alphanumeric methods, and can be used to further harden hard-biometric methods against spoofing. Reid and Nixon [14] attribute this to the use of labels and measurements, which take the appearance of mere human depictions, that people can easily grasp. Samangooei, Guo, Mark, and Nixon [20] define these as semantic annotations. Dantcheva, Velardo, D'Angelo and Dugelay [21] present a novel definition for soft-biometrics based on additional biometric traits. Some of which are: color of clothes, material accessories, body weight, physical and behavioral traits. Jain, Dass, and Nandakumar [22] describe soft-biometrics as being able to provide information about an individual but without the ability to create a full distinction between that individual and others [23].

3. Related Work

Taking a lead from a paper on thermal facial recognition [24] and other works which use *soft biometrics* to augment hard biometric methods [25], [14]. A contemporary approach, to person recognition, using thermal imaging is being proposed. With a focus on facial data, *hard* and *soft biometric* attributes [4]. Most person recognition systems, depending on the operating environment and scenario, have to choose sensor type. Also, based on the sensing setup and output, have to choose a prediction approach. Irrespective of these, the following steps are common to most systems:

- Detection.
- Feature extraction.
- Semantic analysis.
- Feature matching.
- Recognition result.

Most systems begin with detecting a biometric aspect required for recognition such as head, gait, face, finger, and iris. Then separating it from the rest of the image. This step, which is the detection (data acquisition and pre-processing) step, is an integral part of the sensor device setup. It requires a lot of accuracy and is not always discussed. As the performance of the subsequent steps and the system as a whole, depend on it. As a result of the technicalities involved in the first step, the focus of this research project would begin with the second step. Which is about feature extraction. This chapter attempts to present an overview, of the work that has been done in the field of person recognition over the past few years, focusing on the techniques and algorithms used.

3.1. Person Recognition Methods

Experiments were done by, Bhattacharjee, Seal, Ganguly, Nasipuri, and Basu, with the Terravic Facial IR Database. Using Haar wavelet transform and Local Binary Pattern (LBP) [24]. Martinez, Binefa, and Pantic propose recognition, in thermal images, using the detection of the eyes, nostrils and mouth. Also the subsequent decomposition into a feature vector with Haar wavelets, and classification using SVM and Gentle boost [26]. Herrmann, Müller, Willersinn, and Beyerer achieved an error rate reduction, of up to 80% on Long Wave Infrared (LWIR) images, using the AMROS and OTCBVS benchmark datasets. The achieved detection, using Maximally Stable Extremal Regions

(MSER), and classification with a Convolutional Neural Network (CNN) [27]. For deep learning Krizhevsky, Sutskever, and Hinton used non-saturating neurons and a GPU implementation of CNN to classify 1.2 million high-resolution images [28].

Shangfei et al. worked on a Natural Visible and thermal Infrared facial Expression database (NVIE). Using expression recognition and emotion inference for 100 subjects with and without glasses. They analyzed with Principle Component Analysis (PCA), Linear Discriminant Analysis (LDA), and the Active Appearance Model (AAM) [29]. Kelly, Stefanie, and Tom did a study on, alternate access methods using, infrared sensing, electromyography, oculography, and computer vision for individuals with severe motor impairments [30]. Bhowmik et al. researched the advantages, of thermal face over visible face, for recognition, using IRIS and the Terravic Facial IR Database. They isolated the various IR spectrums, showing the Long Wave Infrared (LWIR) band, as that with the highest emissions and suitability for face recognition. They compared this to Short Wave Infrared (SWIR) and Medium Wavelength Infrared (MWIR) [31]. Ayan, Suranjan, Debotosh, Mita and Dipak Kumar proposed the use of Thermal Minutiae Points (TMP), using blood perfusion data as a feature, for thermal face recognition. They experimented with PCA, LDA, and Equinox, for feature extraction and a multi-layer perceptron back propagation feed-forward neural network for feature classification and recognition of thermal infrared face images [32]. Siu-Yeung, Lingyu, and Wen proposed, using the Modified Hausdorff Distance, to measure the similarity between two feature vectors of thermal faces [33]. Ognjen, Riad and Roberto worked on, improving recognition in visible and, thermal images through the detection of prescription glasses [34].

3.1.1. Background subtraction

This person detection method assumes an image to be part of a continuous stream of pictures and attempts to predict the type of background using the frame sequence in the time domain. A comprehensive overview of the method has been done by Piccardi [35]. The method, however, has a problem with, shadow effects, sudden changes in illumination, occlusions and, scenes within which a person remains static for a long time. These problems can be ameliorated by, equipping cameras with, ground analysis capability [36], thermal imaging and stereo capabilities [37].

Here is a list of background subtraction methods:

- Running Gaussian average
- Temporal median filter
- Mixture of Gaussians
- Kernel density estimation (KDE)
- Sequential KD approximation
- Co-occurrence of image variations
- Eigen backgrounds

3.1.2. Silhouette

This method involves creating the silhouette of an image through binarization [25], using its edges and contours, and matching it with prior models in the persons' database [38]. This method works well for detecting people, using facial images showing countenance, not wearing occlusions or having unique behavioral attributes [39]. The detection of partially occluded persons, was tackled by Wu and Nevatia, by breaking down the image into segments they called edgelets. Then the learning of the features, of each segment, by an AdaBoost algorithm. Then compiling prediction results, made for each segment, to ascertain overall detection accuracy [40].

3.1.3. Appearance, look and feel

This method models the image, based on its appearance and attempts to extract features from it, using the look and feel of the many surfaces that make it up. The features are used to train a classifier algorithm, which learns its various attributes and is able, to make predictions on new images based on the feature set. The method can be used, based on the whole image or, by using bits of the image and summing up the learned features.

3.2. The Whole Image Approach

A novel way of achieving this is, by using wavelets, to analyze the whole image, at a pixel level and computing the difference between the different areas. Some known wavelet transform methods are Continuous Wavelet Transform (CWT), Gabor Wavelet Transform (GWT), Discrete Wavelet Transform (DWT), Haar's Wavelet Transform (HWT) and Fast Wavelet Transform (FWT) [41]. The first known one was DWT invented by, a Hungarian mathematician, Alfréd Haar in 1909.

Fourier transform is also useful for this type of analysis. Its weakness is that it doesn't account for the spatial information only the frequency aspects. These approaches are good for, extracting features from, thermal facial images. Bhattacharjee, Seal, Ganguly, Nasipuri, and Basu used the HWT method, with a Local Binary Patterns (LBP) classifier, to process cropped thermal faces. Divided into frequency bands LL1, LH1, HL1, HH1, where LL1 corresponds to, the lowest frequency band and HH1, the highest frequency band [25]. Figure 5 shows a sketch map, of the wavelet bands from a thermal face, after quadratic wavelet decomposition.

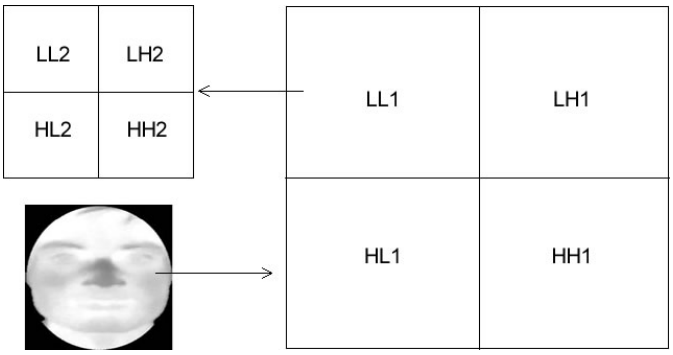


Figure 2: Sketch map of wavelet bands, after quadratic wavelet decomposition of a thermal facial image, showing low and high-frequency areas and sub-bands within the LL1 band [25].

Jones, Viola, and Snow used the wavelet technique, with motion and appearance information, to detect walking people, for outdoor surveillance. After which they classified, them, using a cascaded AdaBoost classifier [42]. For detecting walking people,

Zhu, Yeh, Cheng, and Avidan used the Histogram of Oriented Gradients (HOGs) technique. They applied a cascade of rejectors approach using a weak AdaBoost classifier. Using their method, they were able to reduce computation time [43].

3.3. The Part Image Approach

This approach processes images, by detecting their aspects for person recognition, and decides by combining the various classifications. For full-body images the head, hands, legs, and torso, are detected separately, processed, and classified. Mohan, Papageorgiou, and Poggio worked on detecting people, in cluttered scenes, using an Adaptive Combination of Classifiers (ACC). The method first uses a four-part, detector, separately trained, to detect four segments of the human body, namely: head, legs, right arm, and left arm. After the body parts have been detected, a subsequent classifier is used to classify the results of the four patterns, as a person or nonperson [44].

Wu and Nevatia worked on detecting, partially occluded, humans within video streams. They used part detectors, and weak classifiers, based on features they called edgelets [45]. Their method involved using face, eyes, nose, ears, and mouth separately, before processing and classifying. Trujillo, Olague, Hammoud, and Hernandez experimented with the IRIS data-set. They combined facial feature localization, with the holistic approach using clustering, Eigenface feature extraction and SVM for classification [46].

Martinez, Binefa, and Pantic focused on detecting, facial components (eyes, nostrils, and mouth) in thermal images, using Haar features with GentleBoost algorithm [47]. Al-Khalidi, Saatchi, Burke, and Elphick developed an approach, for tracking a Region of Interest (ROI) on the human face, using thermal video images. The ROI: tip of the nose was tracked using warm and cold facial points of the human face [48].

3.4. History of biometric recognition

Ever since the beginning people have used bodily attributes such as voice, hair, face, and gait, to recognize one another. The idea, that spurred biometrics as a science, was introduced within the field of law enforcement. Its inception was towards the end of the 19th century. When Alphonse Bertillon, police head of the criminal identification department in Paris, introduced the idea of using body measurements, such as skin, eye color, hair, beard, fingers, size, and shape of the head, to identify criminals [11]. This was followed by, a second finding within the same field of law enforcement. The discovery of how human fingerprints are distinct from each other. After which many law-enforcement bureaus began the practice of booking offenders by fingerprinting. The fingerprint data was stored in card files which were replaced by databases. In the hands of law enforcement this method was further improved for determining criminal identities. By using the ability to match fingerprint fragments, from crime scenes, with those in the criminal's database. Beyond law-enforcement biometrics also saw initial application in legal work. Such as in criminal identification, illegal migrant identification, security clearances among employees in sensitive occupations and forensics. Also the identification of ex-convicts, prisoners among suspects, paternal and maternal determinations and much more. In today's world, the private sector, comprising civilians is beginning to include biometrics to alphanumeric methods for person recognition [49].

3.4.2. Verification mode

In *verification mode* a biometric scheme confirms an individual's identity, by comparing the received biometric characteristic, with the owner's biometric template captured in its database. Such systems work by using the assertion of one who is seeking access to perform a one to one comparison and return a true or false. This is done using a Personal Identification Number (PIN) code, an alphanumeric code, a bar code, a login name or smart card.

3.4.3. Identification mode

In *identification mode* the scheme recognizes, a person's identity, by searching the template database for a match. This is done with received biometric characteristic after comparing the input to many records in the database. The check either passes, when a match is found, or declines if the person is not enrolled in the system's database. The *Identification mode* is also used for negative recognition. This entails the system determining that a person is who he or she denies being. This is useful in a situation that a person attempts to use multiple or stolen identities. Negative recognition can however only be used with biometrics. Passwords, keys, tokens, and PINs, can only serve for positive recognition.

3.4.4. A false accept

A *false accept*, also known as a false match, is the confusion of biometric samples from two different users to be, from one person. For a biometric scheme it is represented numerically as the *False Match Rate (FMR).*

3.4.5. A false reject

A *false reject*, also construed as a *false nonmatch*, is the mistaking of two biometric samples from the same user to be from two different individuals. For a system it is represented numerically as the *False Non-Match Rate (FNMR).* Besides recognition error rates: *FMR* and *FNMR*, other parameters which help to sum up a biometric system's performance are the *Failure to Capture (FTC)* and *Failure to Enroll (FTE)* rates [50].

3.4.6. The Failure to Capture rate

FTC rate is the number of times, in percentage, the biometric device is unable to capture a sample when users present their biometric attribute.

3.4.7. The Failure to Enroll rate

FTE rate is the number of occasions, in percentage, users are unable to enlist in the recognition scheme. This could be because of a network issue, a system issue or an issue with the sensor, during biometric characteristic capturing.

3.5. Hard-biometrics

A person's image, normally, comprises both *hard* and *soft biometric* characteristics. The hard-biometric characteristics are -

face, iris, and gait. The *soft biometric* characteristics are - gender, hair, hat, clothes with other occlusions that the individual may have on [51].

3.5.1. Face

This is a commonly used and nonintrusive biometric trait. Its application ranges from authentication in a static fixed background, to one in a dynamic and changing environment. A method used in face recognition involves; 1) the identification of, Regions of Interest (ROI) such as, eyes, eyebrows, nose, lips, and chin and the analysis of their spatial relationship; 2) another is total analysis, of all images that constitute a face, using a weighted aggregate of several model faces, to develop face space and the location of new images within face space. The challenge, with face recognition, is the capturing process.

A fixed and simple background, with known illumination type, is often assumed. However, it is not always the case. A *Facial Recognition System (FRS)* should be able to automatically; 1) Detect the presence of a face within captured images; 2) Locate the face; 3) Recognize the face [52]. The use of a thermal sensor, good feature extraction methods, with deep learning algorithms, would surmount the challenges and increase the capabilities of the system.

3.5.2. Iris

The Iris is the portion of the eye between the pupil and the white of the eye (sclera). Its complex patterns, carry distinct information per individual and, are useful for person recognition. The irises of all persons, even of identical twins, are expected to differ. Iris systems have low *False Accept Rate (FAR)* however their *False Reject Rate (FRR)* can be high [53].

3.5.3. Fingerprint

Fingerprints have been used, for decades, for personal identification. It dates back, to the origin of biometrics, as one of the first methods used for enrolling and documenting human beings. A fingerprint is peculiar for every person. It comprises a unique arrangement of ridges and valleys on the finger surface. For large scale identification, covering millions, an enrolment system comprising multiple fingerprints (all five fingers), from each person, would be more robust than one based on single fingers. However genetic factors, with aging and job-related issues (e.g., motor mechanics, farmers, bricklayers, and manual laborers), affect the general use of fingerprint for automatic identification.

3.5.4. Hand geometry

This recognition approach is based on dimensions of the human palm its length, width, shape, and size. Hand geometry is not affected by manual labor, work factors, or environmental factors, such as heat or humidity. It is however not very distinctive, and in a large population, more than one person could have the same hand geometry. Also factors such as physical impairment, lack of motor function, paralysis and arthritis, affect the universal use of hand geometry for automatic authentication.

3.5.5. Gait

Gait biometrics, is an appearance-based authentication method, which works well with video sequences. The ubiquity of video cameras has made this type of recognition popular. Gait biometrics, is quite distinctive and, can serve for automatic recognition among persons in a large population. It is very useful for recognition over distances, and under poor resolution conditions, but is affected by factors such as fast-changing backgrounds, occlusions, costumes and differences in lighting.

3.5.6. Voice

Voice recognition comprises both hard and soft biometrics. The *hard biometrics* being aspects such as, the mouth, vocal cavity, vocal tracts and lips. While the *soft biometrics* consist of emotional states such as age, health conditions, etc. Voice is not very distinguishing, and would not be useful for automatic recognition, among persons in a large population. A voice recognition scheme may be text-independent or text-dependent. The text-independent types, being more arduous to design than text-dependent ones and, offering better security. Voice is also affected by environmental influence such as noise and echo.

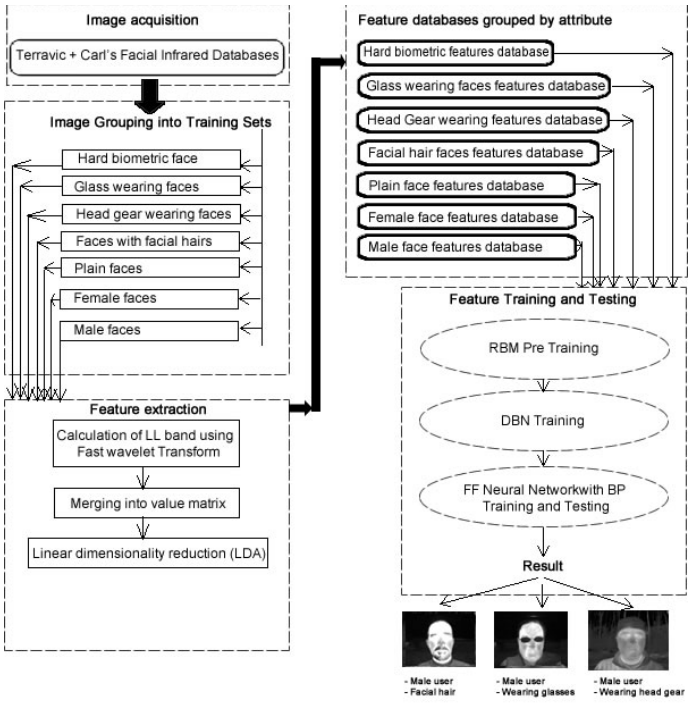


Figure 3: The proposed system, for multi-biometric thermal face recognition, using FWT image compression approach, LDA feature extractor, and RBM, DBN with FFNN classifiers.

3.6. Hybrid systems

These systems merge *soft biometric* data, with *hard biometric* information, to increase overall accuracy. Described by Jain, Dass, and Nandakumar, it suggests the complementing of primary biometric information gathered from fingerprint and face with *soft biometric* data, to achieve resistance to circumvention acceptability, collectability, permanence, universality and distinctiveness [54]. Scheirer, Kumar, Ricanek, Belhumeur and

Boult referred to this as, improving match scores from, face recognition attempts resulting from incomplete measurements done at match time. By using Bayesian Attributes Networks to combine descriptive attributes, such as *soft biometrics,* with primary biometric observations [55]. Also the use of multi-biometrics, comprising soft and hard biometrics, in a recognition system helps in hardening it against unwanted error rates and spoof attacks (such as fraudsters attempting to forge an identity). It also helps, in the case of insufficient population coverage, where little is known about most candidates monitored, using a recognition system [56].

4. The Proposed Method

The method was researched using thermal facial replicas from the Terravic Facial Infrared Database [9] and Carl's database [10]. These were processed using the Large Time-Frequency Analysis Toolbox (LTFAT) in Octave [57] for FWT and Byte fish face recognition toolbox [58] was used for facial image set importation tools. Rasmus Berg Palm's deep learning toolbox [59], was used for, Restricted Boltzmann Machines (RBM), Deep Belief Network (DBN) and Feed Forward Neural Network (FFNN) algorithms. Figure 3 shows a schematic of the proposed system's structure.

4.1. The Terravic Infrared Database

This is one of the two openly usable thermal face image databases. The other is IRIS (Imaging, Robotics and Intelligent System) Thermal/Visible Face Database. It is a subset of the OTCBVS benchmark and is good for experiments in the area of computer vision algorithms. Along with *hard* and *soft biometric* detection using thermal images. The data set comprises thermal sequences of 20 different people, each sequence having variations such as, right, left, front, outdoor, indoor, hat and glasses. All captured, with a Thermal-Eye 2000AS Raytheon L-3 sensor, with a focus on facial analysis for thermal imagery. The images have an 8-bit grayscale with, JPEG format, each one being 320 x 240 pixel in size [8]. Four independent training sets, having *soft biometric* attributes, were established from the dataset. Case study images from each one are presented in Figure 4, Figure 5, Figure 6 and Figure 7.

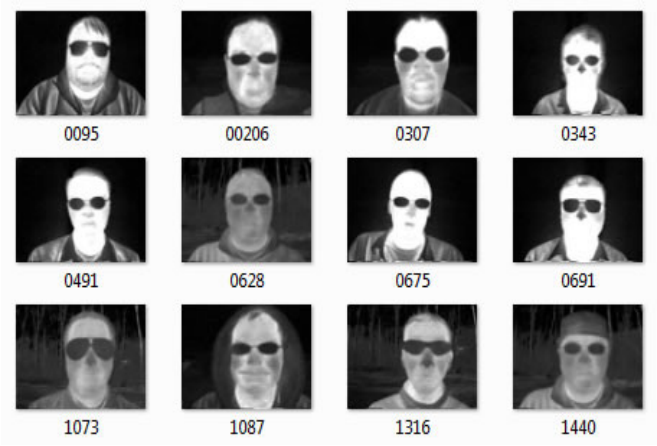


Figure 4: Case study images having glasses-wearing charateristic, from the Terravic Facial Infrared Database.

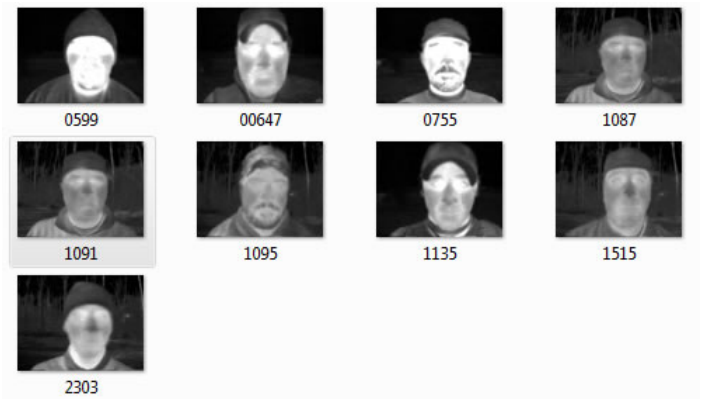


Figure 5: Case study images having headgear charateristic, from the Terravic Facial Infrared Database.



Figure 6: Case study images having facial hair charateristic, from the Terravic Facial Infrared Database.

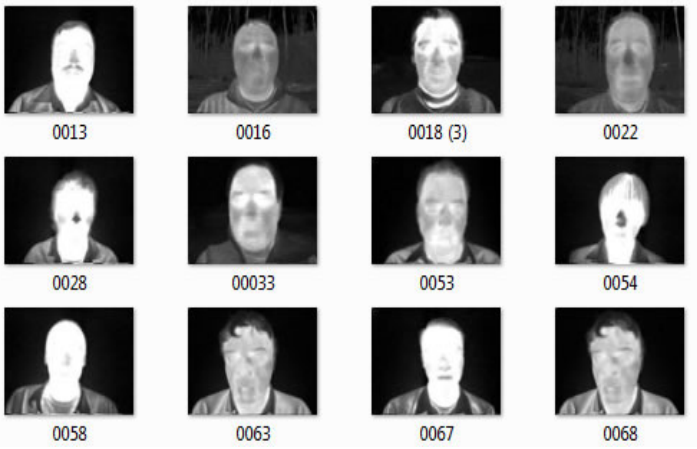


Figure 7: Case study images having plain face charateristic, from the Terravic Facial Infrared Database.

4.2. Fast Wavelet Transform (FWT)

Thermal images, which are a mixture of mixed and relevant detail, can be filtered using a wavelet transform technique. This breaks down the image into regions of high and low-frequency comprising frequency and spatial components, through a compression technique. Graphically these can be represented as sub-bands. Some common wavelet transform approaches are Fast Wavelet Transform, Discrete Wavelet Transform, Gabor Wavelet Transform and Continuous Wavelet Transform [60]. Fast Wavelet Transform (FWT) was used with Octave, in the experiments, because of its application to thermal facial imaging [61]. FWT is a constituent of the Large Time-Frequency Analysis Toolbox (LTFAT) [62].

4.3. Linear Discriminant Analysis (LDA)

LDA accomplishes the derivation of points of highest variance, from images in feature space. It does this by making sure the

within-class distance (S_w) is less, for images in the same class and between class distance (S_b) more, for images in seperate classes. It advances the Principal Component Analysis (PCA) [63] method, through the way it derives the covariance matrices, applying linear discriminants [64].

$$AV = \lambda Y \quad (1)$$

Where $A = Covariance\ matrix, \mathrm{V} =$ Eigen vector matrix,

λ = Diagonal matrix of analogous Eigen values.

$$S_b V = \lambda S_w V \quad (2)$$

Where, $\mathrm{V} =$ The eigenvector matrix, $\lambda = Eigen\ values$.

4.4. Restricted Boltzmann Machines

Deep-belief networks are constituted using simple dual-layer neural networks called RBMs. Their layer design comprises both visible and hidden layers. Having same layer nodes restricted from communicating and different layer nodes allowed to communicate. Each node being a standalone center for data processing. Accepting its input and randomly choosing whether to transmit that input or not. Some applications of the RBM algorithm being regression, feature learning, classification, dimensionality reduction, topic modeling and collaborative filtering [65].

$$\mathrm{f}((\mathrm{x} * \mathrm{w}) + \mathrm{b}) = \mathrm{a} \quad (3)$$

Where, $x = Input$, $w = Weights$, $\mathrm{b} = \mathrm{Bias}$, $\mathrm{f} =$ Activation function, $\mathrm{a} = \mathrm{Output}$.

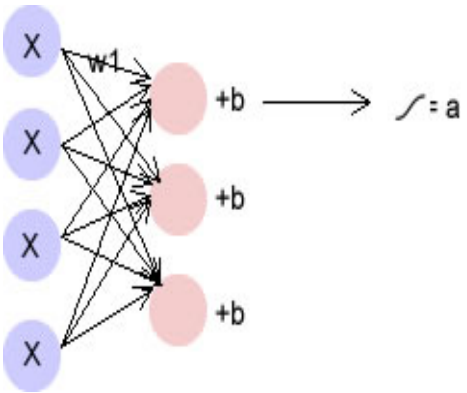


Figure 8: RBM network depicting *inputs x*, *weight w1* for a connected node, bias for hidden layer and outputed activation function.

4.5. Deep Belief Networks (DBN)

When two RBM layers constitute a network, with many hidden layers, they become a Deep Belief Network (DBN). With this structure DBNs are able to learn highly structured data, such as thermal images. The irregular layer design of the RBM, makes its input to propagate through *Contrastive Divergence,* such that good classification is achieved irrespective of network depth. Using an unsupervised algorithm, the RBM is first pre-trained. After which the error received at the output, is sent to the input through backpropagation. After which supervised learning begins, by adjusting the weights on the successive forward pass [66]. The expected outcome is to reconstruct a target vector, using a speculative approach, with as little inaccuracy as possible. This is done, after several forward and back passes. Tools from Rasmus Berg Palm's deep learning toolbox, for RBM and DBN, have been used [59]. The Deep Belief Net design comprises an RBM having 100 hidden layers and a DBN consisting of 100 by 100 RBM units. The RBM algorithm weights were used to commence the Deep Belief Network after pre-training with 10 epochs of 100 batches each. Using different start-up seeds, the DBN algorithm was trained many times, for each concourse. Running 100 batches per epoch, the DBN algorithm completed 10 epochs, for each *soft-biometric* class. After which the number of iterations was plotted against the average reconstruction error. An error of 0.026% was achieved for plain faces, 6.05% for glass faces, 0.026% for head gear faces, and 0.026% for facial hair faces. A feed-forward neural network classifier was initialized using the final DBN weights. Here is the DBN architecture.

Table 1: Deep Belief Network Architecture

Feature	**Description**
Activation function	Sigmoid
Input Layer	1
Output Layer	1
Momentum	0
Alpha	1
Batch scope	10
Epochs	100
RBM start-up weight	100 by 784 zero matrix
RBM unit scope	100 layers
DBN start-up weights	100 by 784 zero matrix
DBN scope	100 by 100 RBM units

4.6. Feed Forward Neural Network

A FFNN has been applied, as the secondary classifier, for the algorithm. After target vector reconstruction, the DBN, was spread out into a FFNN. This was trained by applying 10 epochs of 100 batches respectively and used to classify sample images from the *soft-biometric* groups.

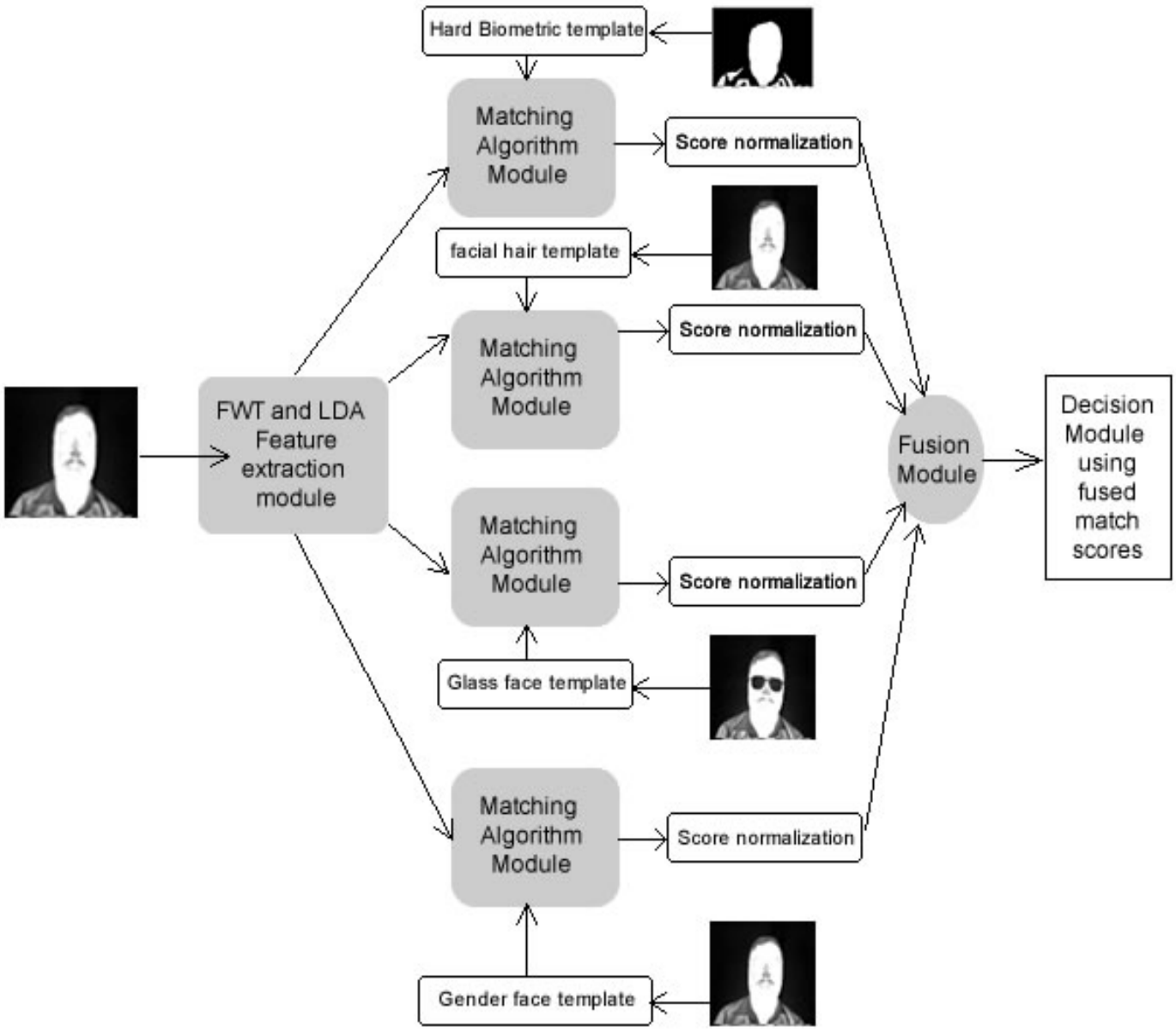


Figure 9: Proposed multi-biometric person recognition system, with single sensor thermal imaging, using a *parallel mode* of operation, with the score fusion approach, for soft and *hard biometric* traits; images from the Terravic Facial Infrared Database.

4.7. Training data acquisition

To design the biometric system, sample images were acquired using public and private thermal facial datasets, namely: the Terravic Facial Infrared Database [9] and Carl's database [10]. The former was used in learning soft and hard-biometric attributes, while the latter for gender attributes, of female and male.

4.8. Image pre-processing and feature extraction

The behavioral and physical attributes of persons, under low lighting conditions can be ascertained, using thermal imaging. Behavioral (outward) attributes such as cap, glasses, hat and facial hair with physical attributes such as head shape and size can be used to classify persons and afterwards recognize them. For thermal imaging a sturdy feature extraction method is requisite. Firstly, Fast Wavelet Transform (FWT), for filtering the noise from the thermal images and secondly, Linear Discriminant Analysis (LDA), for extracting their principal components.

4.9. The multi-biometric recognition method

Using images from the Terravic Facial Infrared Database, the hard-biometric trait (shape of a head) along with *soft-biometric* traits (glasses, a cap, a mustache, beard or gender) will be detected from a single image. Using a single thermal sensor with matching algorithms trained to recognize each feature. After which the outcome of each match would be, fused and, used to make the final decision. This is illustrated in Figure 9.

4.9.1. Multi-biometric recognition modes

To implement multi-biometrics a system can function using any of these three diverse modes of operation, namely: 1) The

serial mode; 2) The *parallel mode* and; 3) The *hierarchical mode* [64].

4.9.1.1. The serial mode

Operating in the *serial mode,* the recognition result gathered from verifying the first trait, is used to reduce the scope of possible matches. The same is done, with subsequent traits, until a final trait is used, to complete the matching process.

4.9.1.2. Parallel

Operating in the *parallel mode*, recognition time is reduced by verifying all the traits concurrently, then using a fusion approach, to sum the matching scores from each search. After which, the result is measured against a threshold score "*s"* before, a final decision is made.

4.9.1.3. Hierarchical

In the *hierarchical mode* of operation, results from respective classifiers, for each biometric evidence, are combined using a staggered, trait by trait, approach with the decision made after a final fusion stage.

4.9.2. Types of multi-biometric fusion

In a *multi-biometric* recognition system's evidence can be combined, before or after the matching process, using different schemes, some of which are: 1) A *sensor level* scheme; 2) A *feature level* scheme; 3) A *match score level* scheme and; 4) A *decision level* scheme [67]. Figure 9 shows a schematic of the proposed multi-biometric system, comprising several matching algorithms and a fusion module before decision, using the fused match scores.

4.9.2.1. The sensor level fusion scheme

This applies to a *multi-sensor* system. It works with *fusion* at the unprocessed data level, using the inputs from each sensor. The received traits are combined to form a single trait, from which features are extracted. An example, for face biometrics, is a system that combine inputs from visible, near-Infrared and thermal spectrum sensors [10].

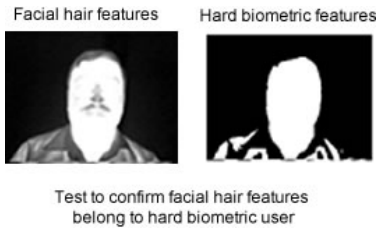
4.9.2.2. The feature level scheme

This method fuses processed features, from each biometric trait, into a final feature set with which the person is recognized. An example is the fusion of facial linear discriminants, from images of the shape of a user's head, with a *soft-biometric* trait such as gender [68]. To form a single feature set with which a user can be recognized.

4.9.2.3. *The match level scheme*

This method sums up, numerical, scores outputted from each classifier. The sum is used, to make the final decision, based on comparison to a threshold score *s*. An example is matching scores from verifying different *soft-biometric* traits such as, particular head shape, the presence of facial hair, glasses or male gender, as shown in Figure 9. Table 2 below illustrates the *match level* scheme with *score fusion*.

Table 2: Matching Algorithm With Score Fusion Method

Matching algorithm with the input image to the left	**Score**
 Test to confirm facial hair features belong to hard biometric user	Score 1, if the hard-biometric features in the input image belong to the authentic user (with facial hair). Score 0 if they do not.
 Test to confirm facial hair features belong to glass face user	Score 1, if the glass wearing features in the input image belong to the authentic user (with facial hair). Score 0 if they do not.
 Test to confirm facial hair features belong to male gender user	Score 1 if the male gender features in the input image belong to the authentic user (with facial hair). Score 0 if they do not.
Total score	3

4.9.2.4. The decision level scheme

This is a voting scheme in which each classifier makes its own decision after verifying the biometric trait presented to it. All the votes are presented at the output. A majority of positive votes indicate an identity as verified. While negative votes indicate an identity as unverified.

4.10. The approach

The single thermal imaging sensor approach, proposed here, will use the *parallel mode* of operation, as shown in Figure 9 with the matching classifier algorithms shown in Figure 3. It will

reduce biometric system errors, by working as a hybrid system in verification mode, using the whole image approach for face biometrics. By having each algorithmic cluster, verify a particular feature trait and, yield a matching score. Then use the *match level* scheme, to make its final decision, through score fusion, as shown in Table 2.

The matching algorithms will each comprise, Fast Wavelet Transform (FWT) and Linear Discriminant Analysis (LDA) for preprocessing, the removal of noise and retrieval of the spatial, time-frequency components. The isolation of the principal components of each image would be done using LDA. The pre-processed feature set will each be classified using, an ensemble of Restricted Boltzmann Machines (RBM) algorithms, Deep Belief Network (DBN) algorithms and Feed Forward Neural Network (FFNN) algorithms. For verifying, each feature trait, and yielding a matching score, for final decision making through *score fusion*.

5. Experiment One: Image Pre-processing and Feature Extraction.

These facial recognition experiments were conducted, for *hard* and *soft-biometric* thermal images, using FWT and Linear Discriminant Analysis (LDA) for feature extraction. FWT is part of the Large Time-Frequency Analysis Toolbox (LTFAT) and is available as a package in GNU Octave.

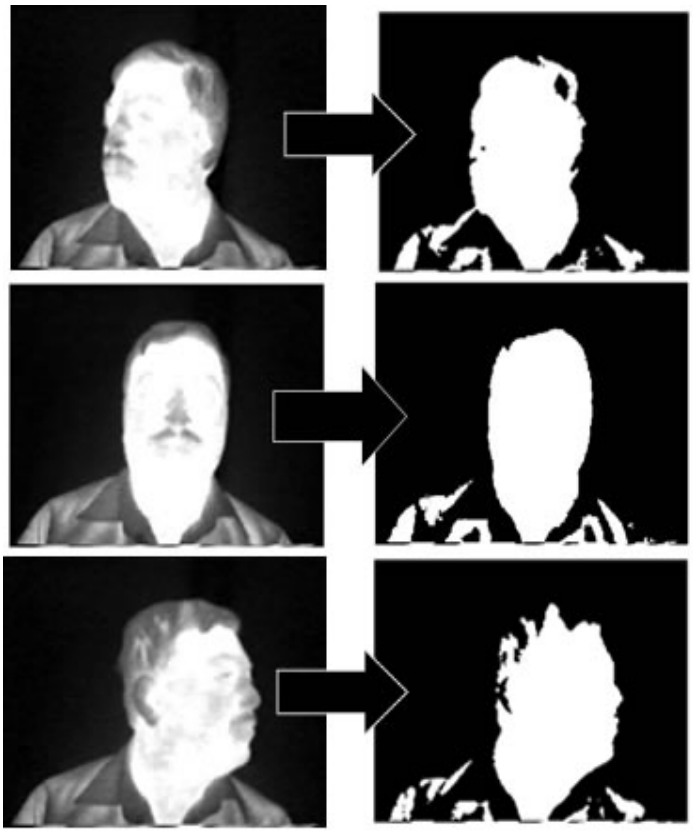


Figure10: Image from the Terravic Facial Infrared Database, before and after binarization.

5.1. Hard-biometrics

The image processing was first done to extract the hard-biometric traits from the facial images. This was achieved by image binarization, using octave, to retrieve the shape of the head from each of the images.

5.1.1. Image binarization

This was done using a threshold technique, to determine the head size and shape. It was applied for extracting the hard-biometric attribute of the thermal images. The technique is a common pre-processor used in OCR (Optical Character Recognition), and is useful in pattern recognition, for exposing the uniqueness of an image.



Figure 11: Multiple exposures of the same person, from different angles, before binarization and LDA, images from the Terravic Facial Infrared Database.

5.1.2. Hard-biometric FWT Experiments

The hard biometric features from the thermal facial images, shown in Figure 11, after binarization were extracted using FWT. Figure 12 shows the wavelet sub-bands. The entire image set is a matrix of size 76800 by 144. This was processed to yield four sub-bands after FWT. The first band, is a matrix of size 9600 by 144, the second a matrix of size 19200 by 144, the third a matrix of size 19200 by 144, and the fourth a matrix of size 38400 by 144. The four sub-bands were merged to form a value matrix, rich in the spatial and frequency components, which served as input to the LDA stage.

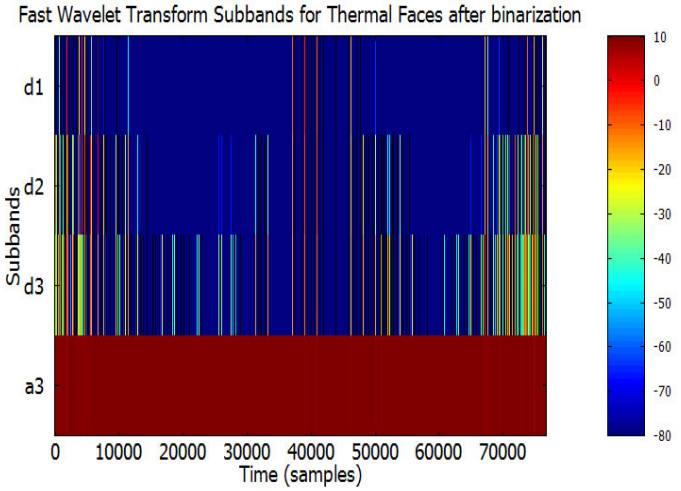


Figure 12: FWT of binarized images, of the same person in Figure 11

5.1.3. Hard-biometric LDA Experiments

LDA was performed on the spatial time-frequency features which formed the confidence matrix of wavelet sub-bands. To further extract points of highest variance. Figure 13 and 14 show a visualization of the time-frequency features, before and after Linear Discriminant Analysis, for the hard-biometric features.

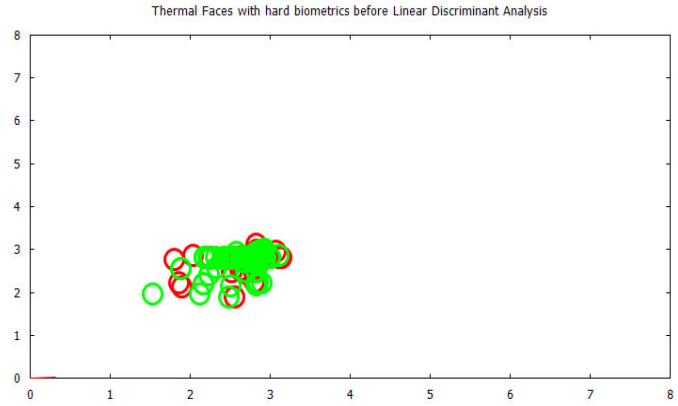


Figure13: Graphical representation of spatial frequency components, of hard biometric features, for thermal faces shown in Figure 10, before LDA

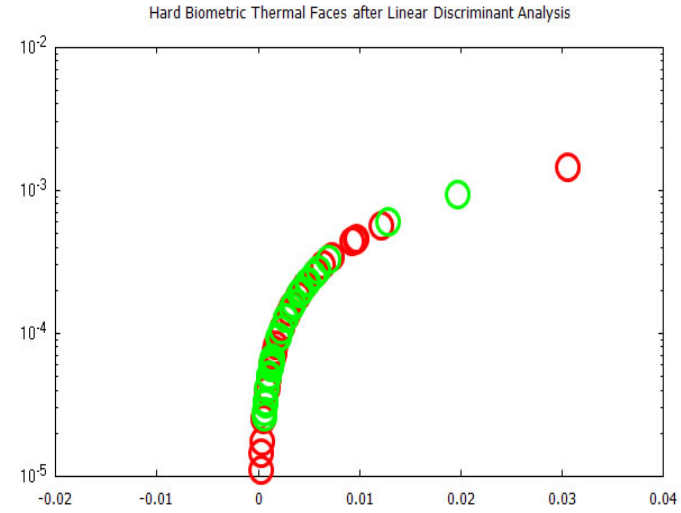


Figure 14: Graphical representation of spatial frequency components, of hard biometric features, for thermal faces shown in Figure 10, after LDA

5.2. Soft-biometrics

This part of the image processing was used for extracting time-frequency and principal components, from each *soft-biometric* attribute. This was achieved by grouping images from the datasets based on the *soft-biometric* attributes presented in Figure 4, Figure 5, Figure 6 and Figure 7.

5.2.1. Soft-biometric FWT Experiments

Figure 15, Figure 16, Figure 17, Figure 18, Figure 19, and Figure 20 show the wavelet sub-bands, for the *soft biometric* attributes after FWT, namely: face wearing glasses, face wearing headgear, face having facial hair, plain face, female face, and male face respectively.

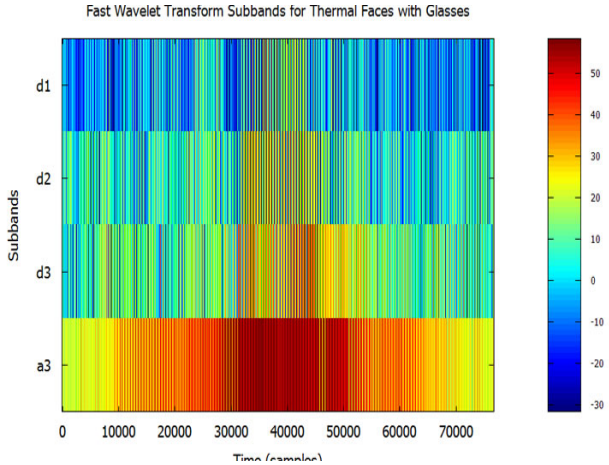


Figure 15: Visualized FWT of face images in glasses, from the Terravic Facial Infrared Database, depicting spatial and frequency information displayed using 4 sub-band

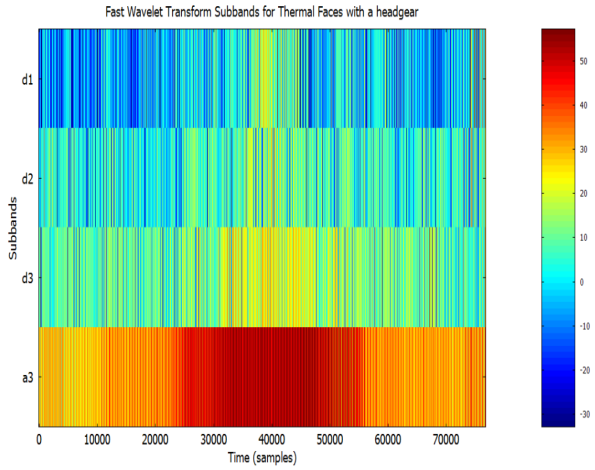


Figure 16: Visualized FWT of face images in headgear, from the Terravic Facial Infrared Database, depicting spatial and frequency information displayed using 4 sub-bands.

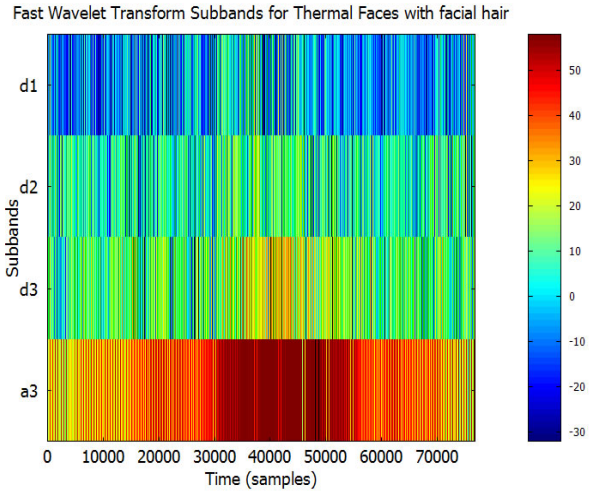


Figure 17: Visualized FWT of face images with facial hair, from the Terravic Facial Infrared Database, depicting spatial and frequency information displayed using 4 sub-bands.

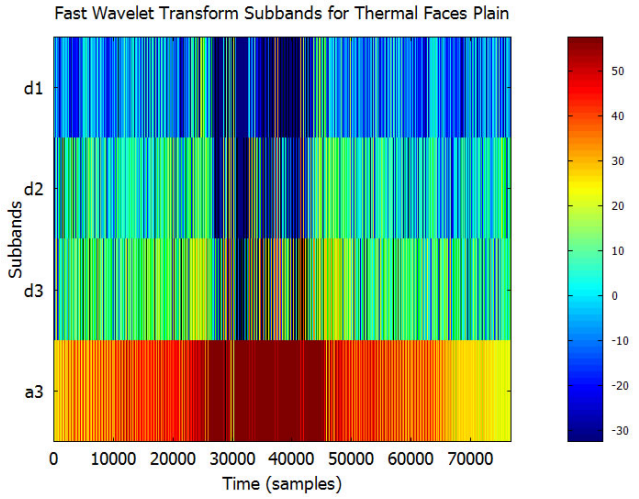


Figure 18: Visualized FWT of plain face images, from the Terravic Facial Infrared Database, depicting spatial and frequency information displayed using 4 sub-bands.

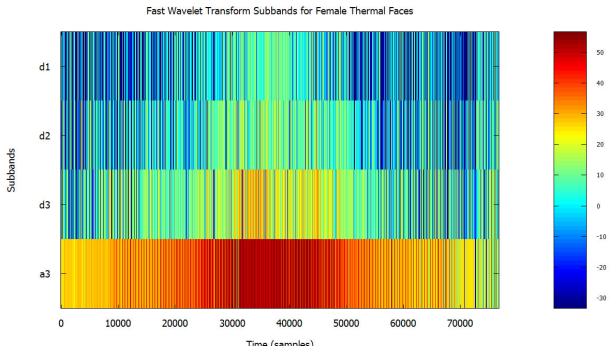


Figure.19: Visualized FWT of female facial images, from the Carl's Infrared Database, depicting time and frequency information displayed using 4 sub-bands.

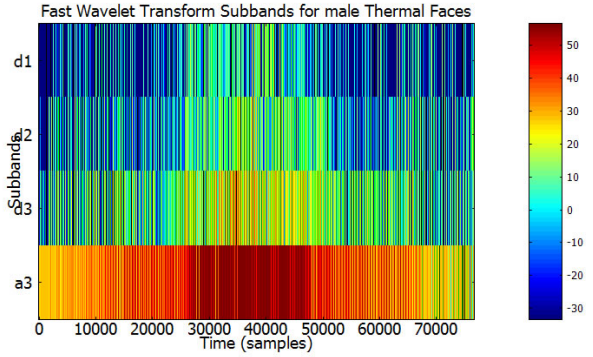


Figure 20: Visualized FWT of male facial images, from the Carl's Infrared Database, depicting time and frequency information displayed using 4 sub-bands.

5.2.2. *Soft-biometric LDA experiments*

After FWT, LDA was performed on the spatial time-frequency features, to derive points of most variance from the confidence matrix, for each *soft biometric* group. The following figures show a visualization of the spatial frequency features before and after Linear Discriminant Analysis, for face wearing glass, face, face wearing cap, face with hair, plain face, female face, and male face.

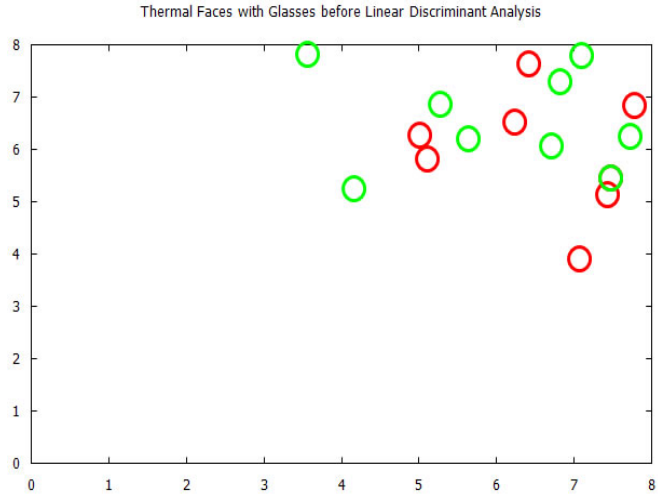


Figure 21: Visualized time-frequency (TF) features, before Linear Discriminant Analysis, for the Terravic Facial Infrared Database images, glass wearing faces

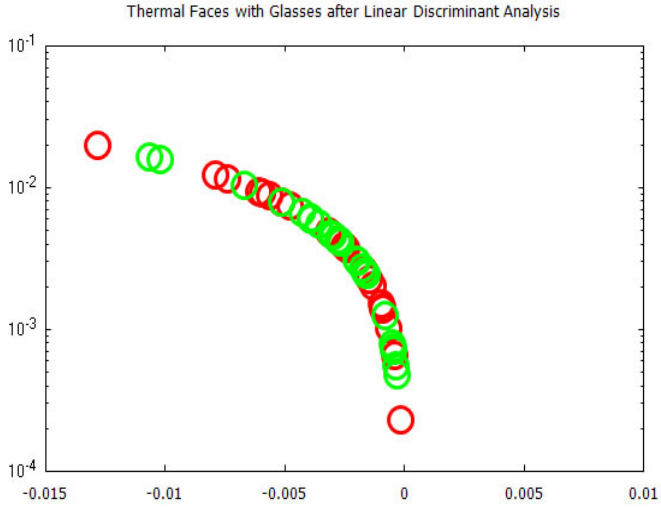


Figure 22: Visualized time-frequency (TF) features after Linear Discriminant Analysis, for images in the Terravic Facial Infrared Database, glass wearing faces

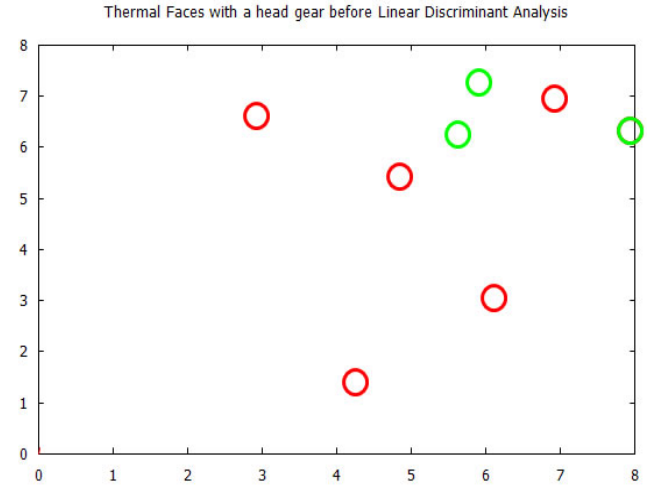


Figure 23. Visualized time-frequency (TF) features, before Linear Discriminant Analysis, for images in the Terravic Facial Infrared Database, cap wearing faces

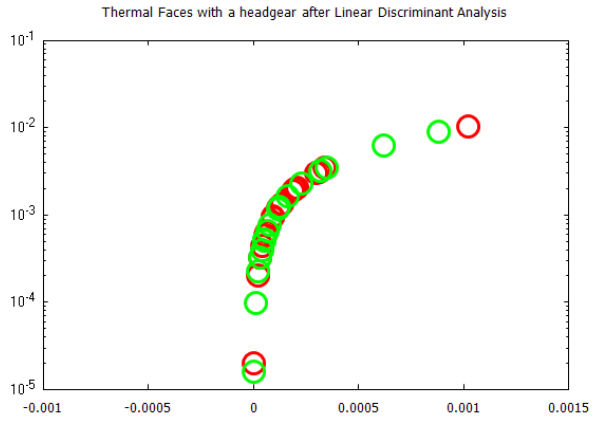


Figure 24: Visualized time-frequency (TF) features after Linear Discriminant Analysis, for the Terravic Facial Infrared Database images, cap wearing faces.

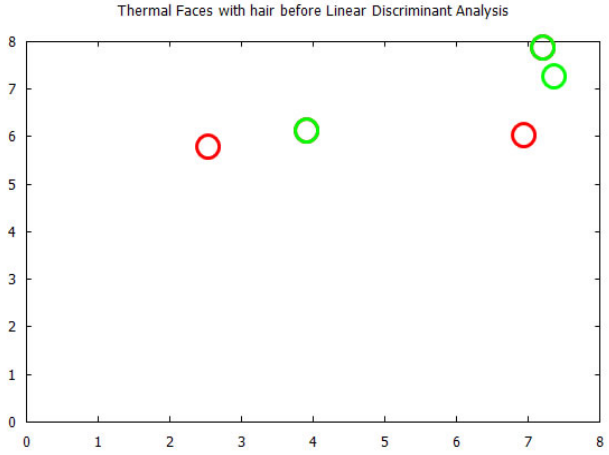


Figure 25: Visualized time-frequency (TF) features, before Linear Discriminant Analysis, for the Terravic Facial Infrared Database images, for faces with facial hair

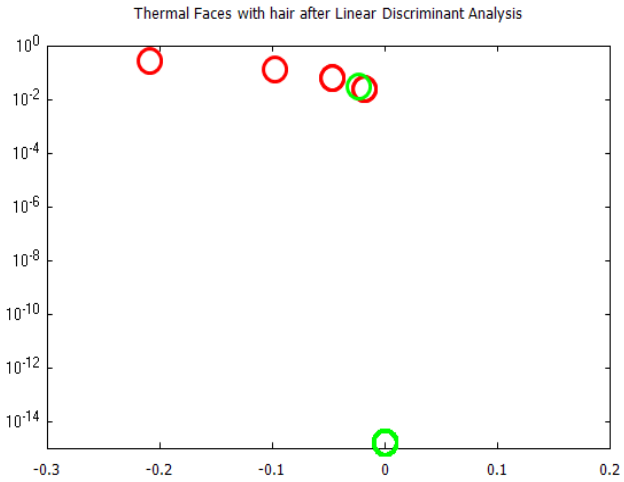


Figure 26: Visualized time-frequency (TF) features, after Linear Discriminant Analysis, for the Terravic Facial Infrared Database images, for faces with facial hair.

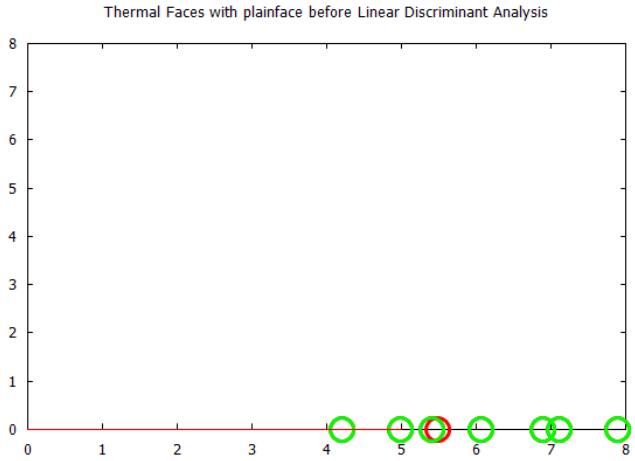


Figure 27: Visualized time-frequency (TF) features, before Linear Discriminant Analysis, for the Terravic Facial Infrared Database images, plain faces

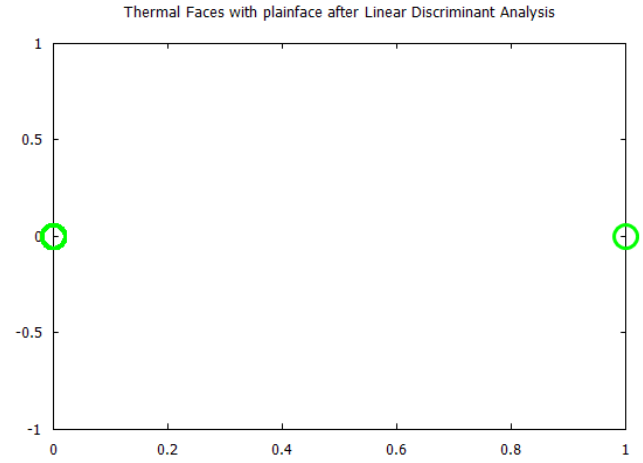


Figure 28: Visualized time-frequency (TF) features, after Linear Discriminant Analysis, for the Terravic Facial Infrared Database images, plain faces

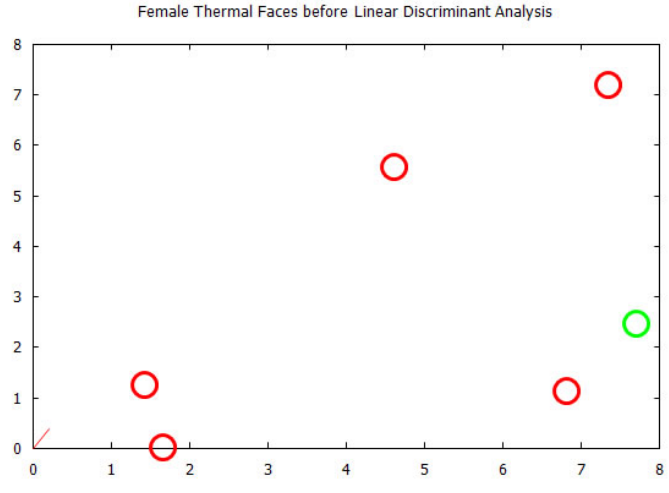


Figure 29: Visualized time-frequency (TF) features, before Linear Discriminant Analysis, for images in the Carl's Infrared Database, female faces.

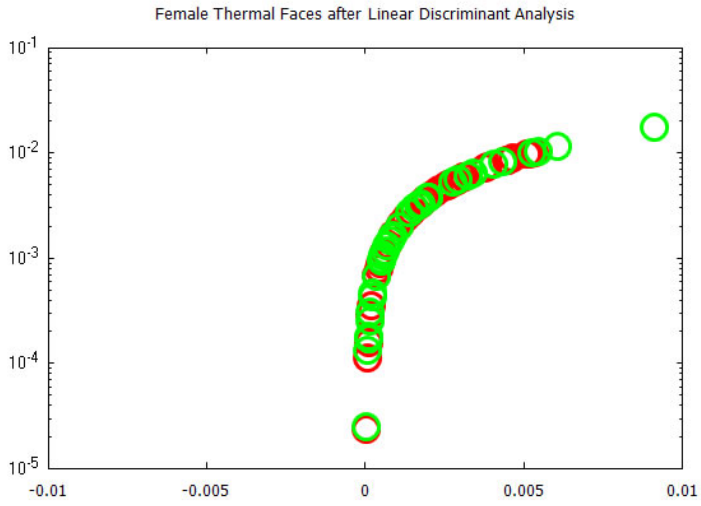


Figure 30: Visualized time-frequency (TF) features, after Linear Discriminant Analysis, for images in the Carl's Infrared Database, male faces.

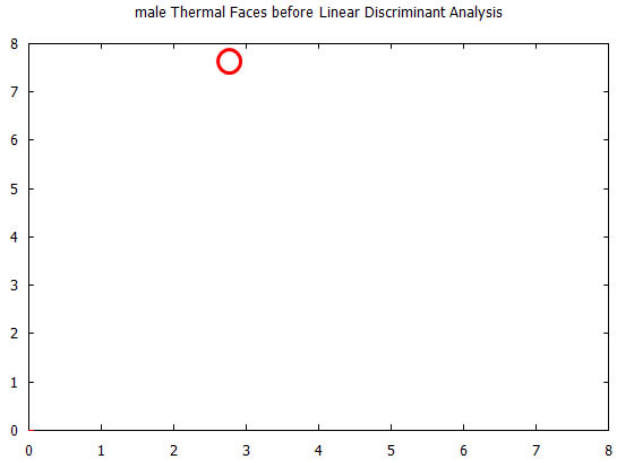


Figure 31: Visualized time-frequency (TF) features, before Linear Discriminant Analysis, for images in the Carl's Infrared Database, male faces

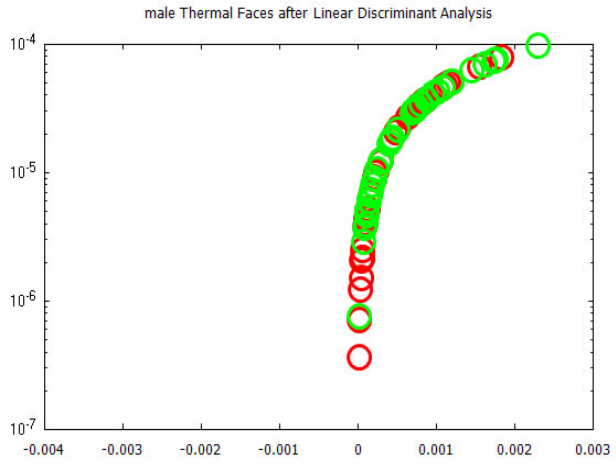


Figure 32: Visualized time-frequency (TF) features, after Linear Discriminant Analysis, for images in the Carl's Infrared Database, male faces

5.3. Summary

The outcome of the feature extraction experiments was stored in a data file using the Matlab format. This was done for both hard biometric and *soft biometric* feature sets. These data files, namely: *Linear Discriminants* for hard biometric features and *Linear Discriminants* for *soft biometric* features. The latter comprising, glass wearing faces, cap wearing face, faces with facial hair, plain faces female faces and male faces. With these the distinguishing features, which describe each of the biometric groups, were known. The objective being to train seven different neural networks, to learn each of the features, and subsequently validate a user's identity using at least two combined. Such that if a user appears, before a thermal sensor, seeking access, without any of the attributes, access would not be granted. But when he or she appears with facial hair, wearing glasses or wearing a cap, access would be granted.

6. Experiment Two: Deep Learning Neural Networks

These are a continuation of the pre-processing experiments reported. They were done, using Rasmus Berg Palm's deep learning toolbox, for RBM, DBN, and FFNN.

6.1. Classification

To achieve target vector reconstruction the *Linear Discriminants* realized from the first experiments were used for classification after training. After pre-training the DBN with RBM and further training and classifying with FFNN, a global minimum error was achieved on each training set group. These were: 0.02887 for hard biometric face, 0.038695 for glass wearing face, 0.02381 for headgear wearing face, 0.024629 for facial hair face, 0.0268 for plain face, 0.02369 for female face and 0.03 male face. Also introducing new test images randomly to the algorithm, while matching the FFNN global minimum error with these known values, demonstrated the *soft-biometric* person recognition approach. Making it possible to identify thermal facial images as: face in glasses, hairy face, plain face, face in a headgear. The following traces show the the full batch classification error, from the RBM, DBN, and FFNN algorithms, for each *soft-biometric* group.

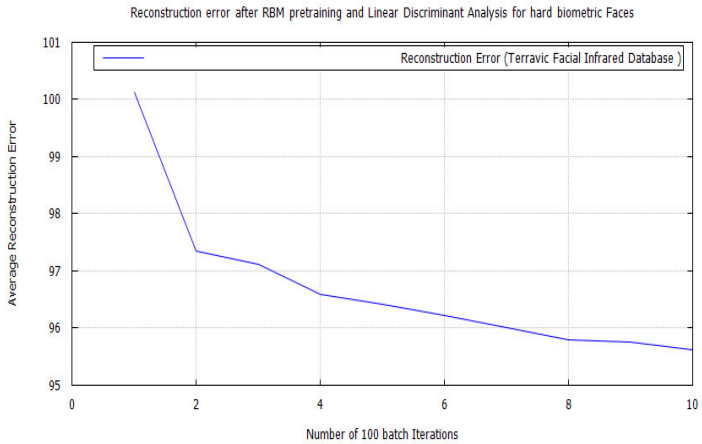


Figure 33: RBM classification error after LDA feature extraction applying the time-frequency features of images from the Terravic Facial Infrared Database, facial hard biometrics user.

6.2. Plots

The plots from each training session, for RBM, DBN, and FFNN, are graphically shown below for each of the seven semantic categories. In the following order: hard biometric face, faces wearing glasses, faces wearing a cap, faces having facial hair, faces plain, female faces, and male faces. Each one shows an improvement, in the algorithm's training error, up until the FFNN session. At which point the algorithm, fully learns the data and makes predictions on it, with little error.

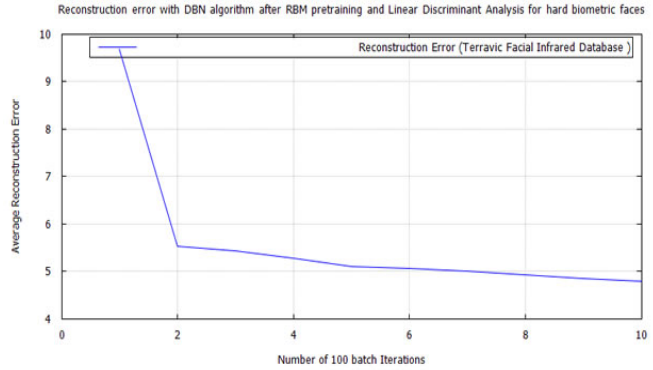


Figure 34: DBN classification error, after RBM pre-training for Terravic Facial Infrared Database images, user facial hard biometrics.

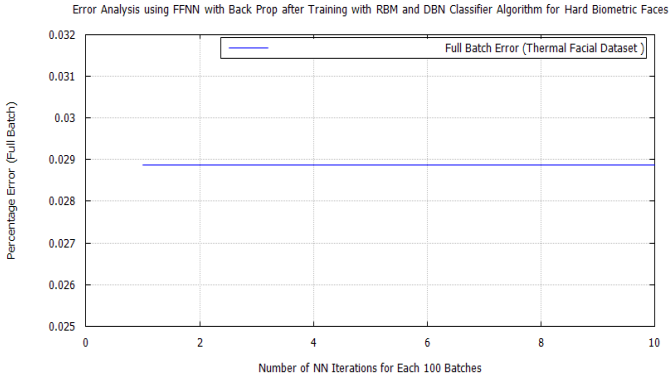


Figure 35: FFNN classification full batch error after DBN target vector reconstruction applying RBM pre-training and LDA feature extraction using the time-frequency features of Terravic Facial Infrared Database images, user facial hard biometric.

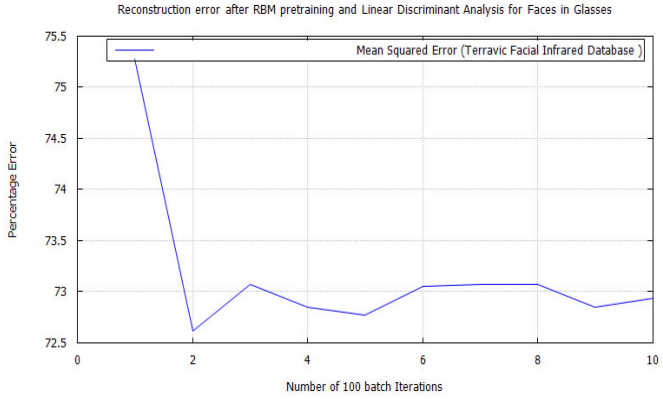


Figure 36: RBM classification error after LDA feature extraction applying the time-frequency features of Terravic Facial Infrared Database images, faces in glasses.

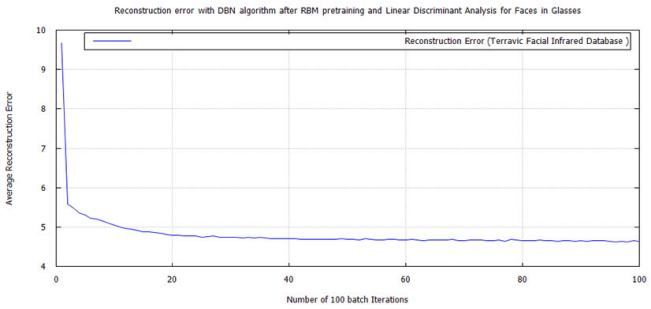


Figure 37: DBN classification error, after RBM pre-training for Terravic Facial Infrared Database images, glass wearing faces.

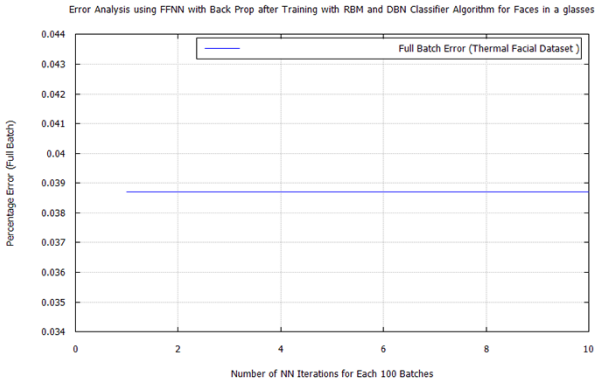


Figure 38: FFNN classification error after DBN target vector reconstruction with RBM pre-training and LDA feature extraction applying the time-frequency features of Terravic Facial Infrared Database images, glass wearing face group.

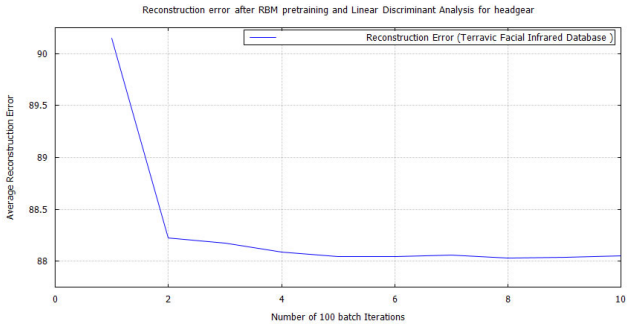


Figure 39: RBM classification error after LDA feature extraction applying the time-frequency features of Terravic Facial Infrared Database images, headgear faces.

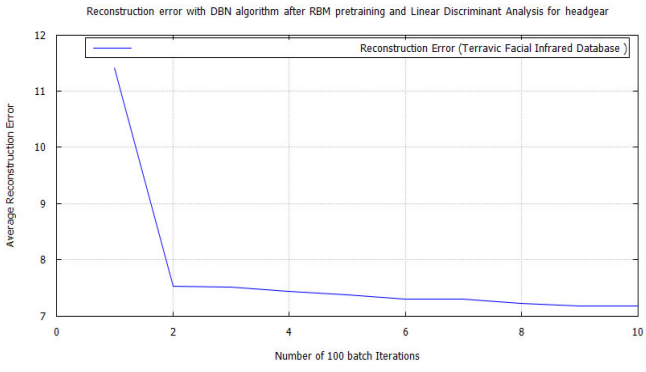


Figure 40: DBN classification error, after RBM pre-training for Terravic Facial Infrared Database images, headgear wearing faces.

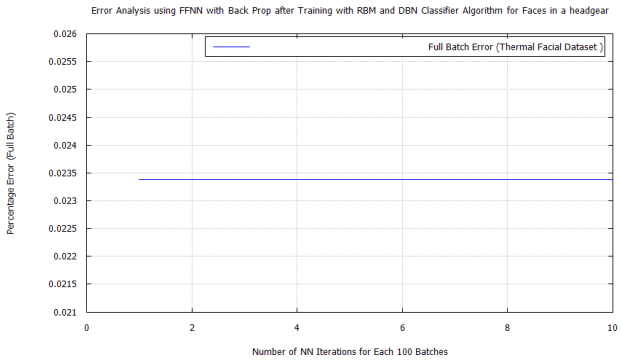


Figure 41:FFNN classification error after DBN target vector reconstruction with RBM pre-training and LDA feature extraction applying the time-frequency features of Terravic Facial Infrared Database images, headgear wearing group.

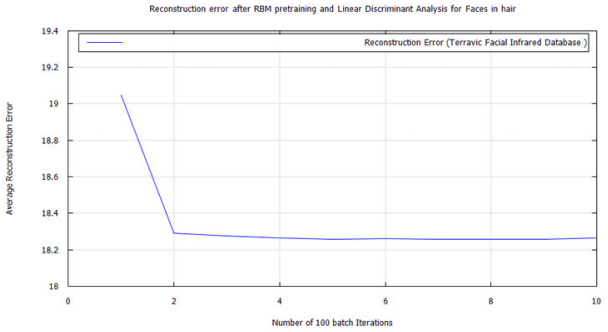


Figure 42: DBN classification error, after RBM pre-training for Terravic Facial Infrared Database images, faces having facial hair.

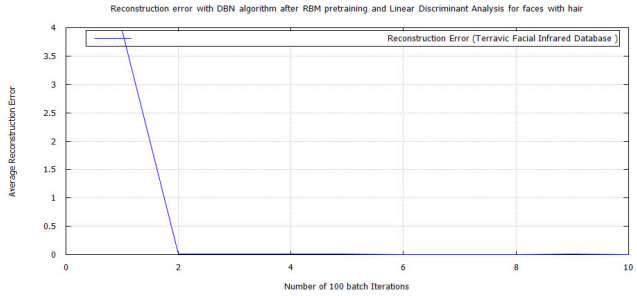


Figure 43: DBN classification error, after RBM pre-training for Terravic Facial Infrared Database images, faces having facial hair.

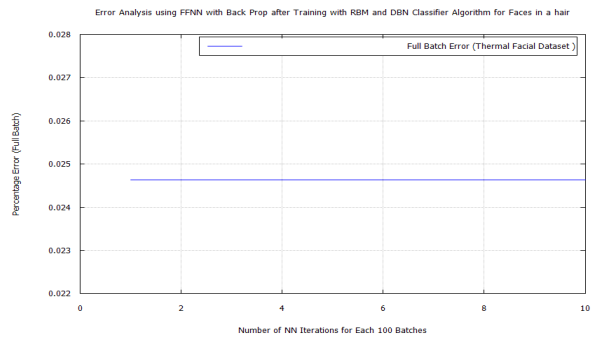


Figure 44: FFNN classification error after DBN target vector reconstruction with RBM pre-training and LDA feature extraction applying the time-frequency(TF) features of Terravic Facial Infrared Database images, hairy faces group.

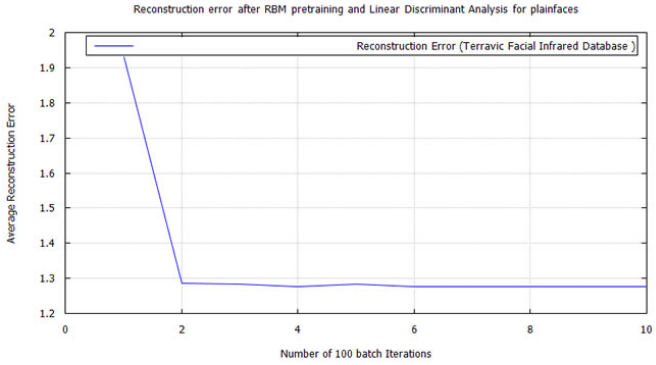


Figure 45: DBN classification error, after RBM pre-training for Terravic Facial Infrared Database images, plain faces.

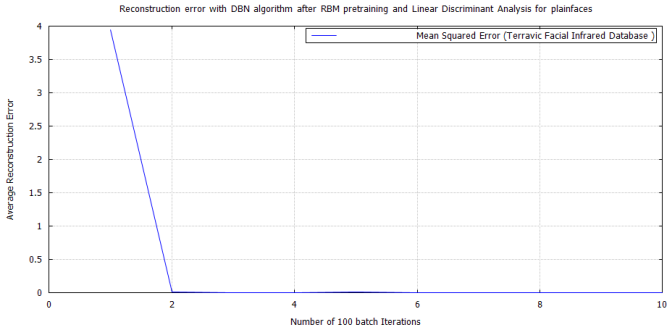


Figure 46: DBN classification error, after RBM pre-training for Terravic Facial Infrared Database images, plain faces.

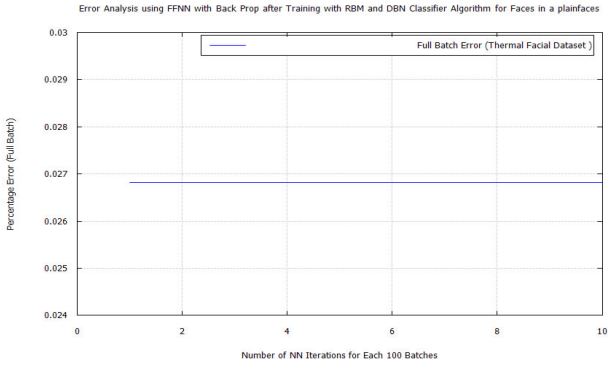


Figure 47: FFNN classification error after DBN target vector reconstruction with RBM pre-training and LDA feature extraction applying the time-frequency features of Terravic Facial Infrared Database images, plain face group.

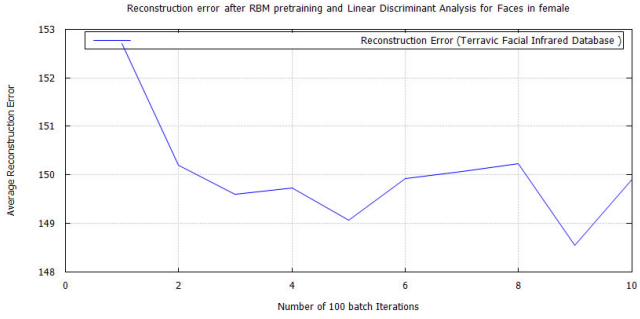


Figure 48: RBM classification error after LDA feature extraction applying the time-frequency features of Terravic Facial Infrared Database images, female faces.

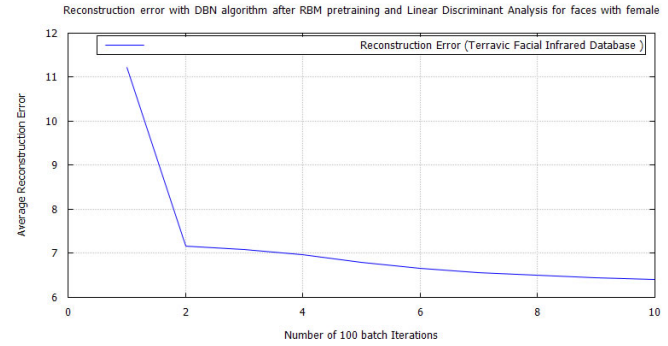


Figure 49: DBN classification error, after RBM pre-training for Terravic Facial Infrared Database images, female faces.

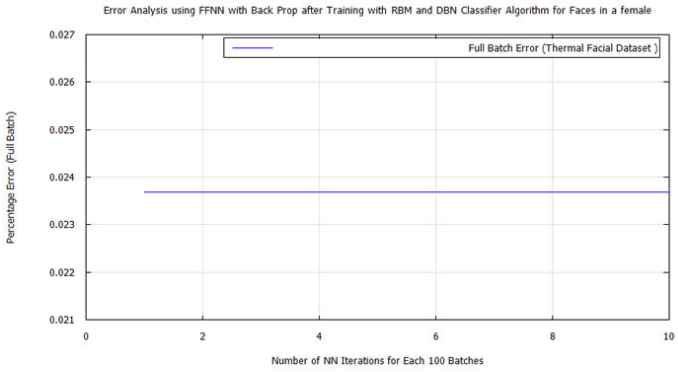


Figure 50: FFNN classification error after DBN target vector reconstruction with RBM pre-training and LDA feature extraction applying the time-frequency features of Terravic Facial Infrared Database images, female face.

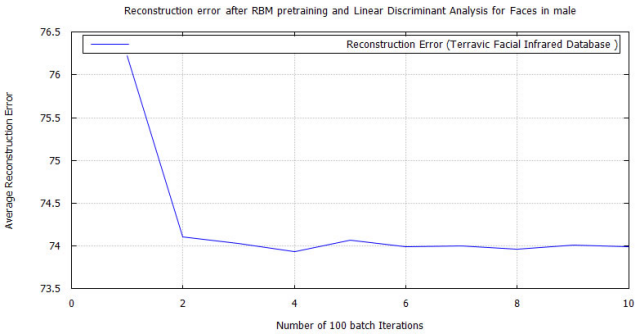


Figure 51: RBM classification error after LDA feature extraction applying the time-frequency features of Terravic Facial Infrared Database images, male faces.

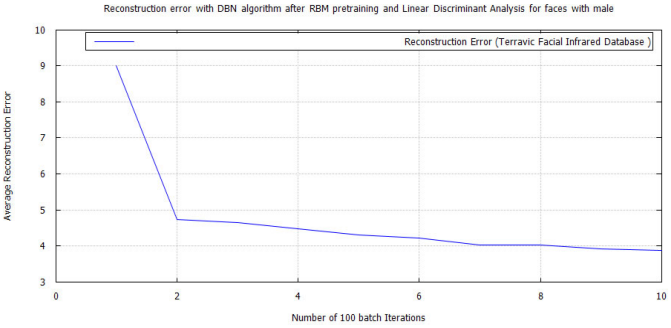


Figure 52: DBN classification error, after RBM pre-training for Terravic Facial Infrared Database images, male faces.

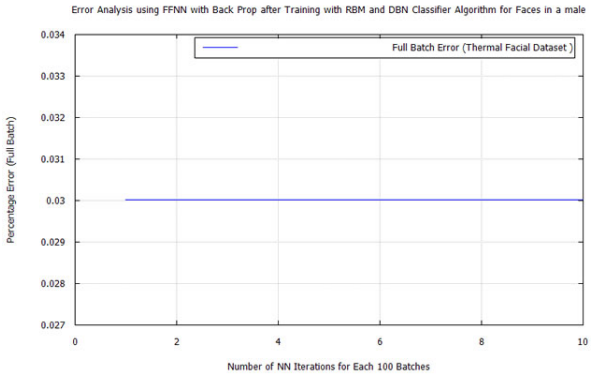


Figure 53: FFNN classification error after DBN target vector reconstruction with RBM pre-training and LDA feature extraction applying the time-frequency features of Terravic Facial Infrared Database images, male face.

Table 3: RBM Test Error Results Per Semantics Attribute Group

Semantic Attribute	**RBM error**
hard biometric face	95.609
Faces in glasses	72.93
Faces wearing a cap/headgear	88.051
Faces with facial hair	18.265
Faces with a plain face	1.2756
Female faces	147.9
Male faces	73.98

Table 4: Test Error Results Per Semantic Attribute Group

Semantic Attribute	**DBN error**
hard biometric face	4.7895
Faces in glasses	4.6364
Faces wearing a cap/headgear	7.182
Faces with facial hair	0.026
Faces with a plain face	0.026
Female faces	6.406
Male faces	3.926

Table 5: Test Error Results Per Semantic Attribute Group

Semantic Attribute	**FFNN error**
hard biometric face	0.02887
Faces in glasses	0.038695
Faces wearing a cap/headgear	0.02381
Faces with facial hair	0.02462
Faces with a plain face	0.0268
Female faces	0.02369
Male faces	0.03

6.3. Result analysis

Applying the RBM pre-trained DBN, with FFNN classifier algorithm method, to the *Linear Discriminants* from the thermal facial images, it was observed that:

- Time-frequency (TF) components, gleaned from Fast Wavelet Transform, aided LDA in the facial thermal image investigation.
- On each of the biometric training sets a low reconstruction error was achieved, using DBN after pre-training with RBM,

through target vector reconstruction. This excepted faces wearing glasses, which needed to be trained using a bigger batch size of 100 to achieve a test error of 4.6364%. This value was further reduced to 0.038695% after additional training with the FFNN algorithm.

- The other biometric groups didn't require additional training and were trained using a batch size of 10.
- With the use of back propagation, DBN pre-trained with RBM algorithm successfully learned the features of the *soft-biometric* training sets. This was possible without lengthy training sessions.
- As in the case of glass wearing face including feature learning, and training with FFNN, aided the further reduction of the classification error.
- The reduced classification error realized, by applying FFNN with DBN pre-trained using RBM and LDA, showed that the algorithm would recognize each *soft biometric* trait and differentiate one from the other.

Table 6: FWT, LDA, RBM, DBN, and FFNN Multi-Biometric Experiment Results Per Semantic Attribute Group

Experiment	**Result**
Feature Extraction FWT	Numbers
FWT Sub-band one	9600 features
FWT Sub-band two	9600 features
FWT Sub-band three	19200 features
FWT Sub-band four	38400 features
FWT Sub-band four + LDA+RBM Hard-biometric features	4.7895 %
FWT Sub-band four + LDA+RBM Glass faces	4.6364%
FWT Sub-band four + LDA+RBM Head gear faces	0.026 %
FWT Sub-band four + LDA+RBM Facial hair faces	0.026 %
FWT Sub-band four + LDA+RBM Plain faces	0.026 %
FWT Sub-band four + LDA+RBM Female faces	6.406 %
FWT Sub-band four + LDA+RBM Male faces	3.926 %
FWT Sub-band four + LDA+RBM+DBN+FFNN Hard-biometric features	0.02887
FWT Sub-band four + LDA+RBM+DBN+FFNN glass wearing faces	0.038695 %
FWT Sub-band four + LDA+RBM+DBN+FFNN head gear wearing faces	0.02381 %
FWT Sub-band four + LDA+RBM+DBN+FFNN faces with facial hair	0.024629 %
FWT Sub-band four + LDA+RBM+DBN+FFNN plain faces	0.0268 %
FWT Sub-band four + LDA+RBM+DBN+FFNN Female faces	0.02369 %
FWT Sub-band four + LDA+RBM+DBN+FFNN Male faces	0.03 %

6.4. Observations

Classifying using FWT, LDA and RBM pre-trained DBN with FFNN algorithms, yielded less error of 0.038695% on the glass wearing semantic group. This was for the following reasons:

- The DBN classifier runs faster and uses fewer system resources because of the structure of its RBM unit. Thus it learns more about the problem within a short period.
- The DBN classifier uses back propagation during training for error reconstruction.
- The DBN classifier, is ensembled with a Feed-Forward Neural Network (FFNN) which further, learns the problem and reduces the classification errors.

Using images in the Terravic Facial Infrared and Carl's databases *Multi-biometric* person recognition in thermal images is demonstrated. This was done using FFNN classifiers, after feature extraction with FWT and LDA, and RBM pre-trained DBN for feature learning. *Hard biometric* features, coupled with outward attributes, such as face wearing glass, face wearing cap, face with facial hair, plain face, female face, and male face, of persons in the datasets, were learned for multi-biometric grouping. Trial images not in the training sets were afterwards recognized multi-biometrically, with both their hard biometric traits coupled with any of the six *soft biometric* traits, using the technique. Table 6 shows a tabulation of the experiment results.

6.5. Summary

The outcome, of the feature classification experiments are stored in a data file using the Matlab format. This is done for both, hard biometric and soft biometric, feature sets. These data files contain the error on the test set, for the prediction of the unknown variable (y) given the known variable (X) after training, for the hard biometric and *soft biometric* features. The latter comprising, faces wearing glass, faces wearing cap, face having facial hair, plain face, female face, and male faces. With these the extent to which the neural network knows each of the biometric, features are noted. Based on the low classification error on each of the test sets, as shown in Table 6, the system can recognize biometric features not in the training set, such as:

- Facial, hard biometric, features of the user used in training the algorithm, as shown in Figure 10.
- Faces of a user when in glasses, as shown in Figure 4.
- Faces of a user when wearing a cap, as shown in Figure 5.
- Faces of a user when having facial hair, as shown in Figure 6.
- Faces of a user when plain without facial hair, as shown in Figure 7.
- The face of a female user.
- The face of a male user.

With these biometric traits additional layers of security can be introduced into a person recognition system. Asides using a user's facial hard biometric features for authentication, two, three, four or five other biometric attributes can be used to make a system more difficult to fool. As a result of such complexity, a forger would find it hard to guess what traits the system would need for recognition to take place. Table 7 shows fifteen possible multi-biometric feature configurations, based on the traits used in this

research, along with their security levels. Each combination can work, using all traits or only one, before a person's identity is confirmed. In either mode, if a user appears before the sensor without at least one requisite *soft biometric* trait based on system requirements, the user's identity would not be verified.

The advantage of the system being that it can acquire all the needed biometric traits using a single sensor. Thus a hacker who watches while a user's identity is being verified, using this method, would be unable to know exactly what biometric attribute to forge.

Table 7: Multi Biometric Feature Configurations and Security Levels

	Biometric Features	**Security Level**
1	Countenance + Glasses	Robust
2	Countenance + Cap	Robust
3	Countenance + Facial hair	Robust
4	Countenance + Plain face	Robust
5	Countenance + Gender	Robust
6	Countenance + Glasses + Cap	More Robust
7	Countenance + Glasses + Facial hair	More Robust
8	Countenance + Glasses + Plain face	More Robust
9	Countenance + Glasses + Gender	More Robust
10	Countenance + Glasses + Cap + Facial hair	Much More Robust
11	Countenance + Glasses + Cap + Plain face	Much More Robust
12	Countenance + Glasses + Cap + Gender	Much More Robust
13	Countenance + Glasses + Facial hair + Gender	Much More Robust
14	Countenance + Glasses + Cap + Facial hair + Gender	Most Robust
15	Countenance + Glasses + Cap + Plain face + Gender	Most Robust

7. Experiment Three: Multi biometric person recognition

Using the trained algorithm cluster, with test results tabulated in Table 12. The system was tested in *verification mode* with the person in Figure 10 as a user. Working with a multi-biometric combinations of traits: Countenance + Glasses + Facial hair + Gender. Which are the user's hard-biometric features, glass wearing features, facial hair features and gender traits. The hard-biometric trait is taken as X (features to be learned) while the facial hair trait as the base for comparison y (unknown feature to be predicted y). Tests are done to confirm the other three soft-biometric characteristics and verify his identity.

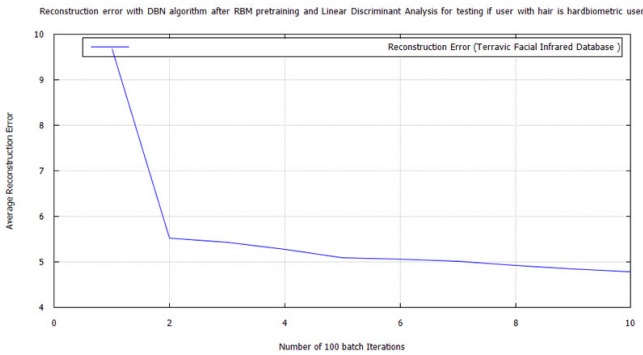


Figure 54: Error for facial hair algorithm testing hard biometric features for the presence of facial hair using RBM, LDA, and FWT for images from the Terravic Facial Infrared Database

7.1. Tests1 hard-biometric and facial hair algorithms

The first test is conducted to verify the hard-biometric features of the user for the presence of facial hair. By using the X variables after LDA, of the hard-biometric features, and the y variables after LDA, of the facial hair features. The RBM, DBN and FFNN tests were conducted to verify, that the user with the hard-biometric features (presented to the sensor), has facial hair as the authentic user. The FFNN result is shown, in comparison with those from the training test, in Table 8.

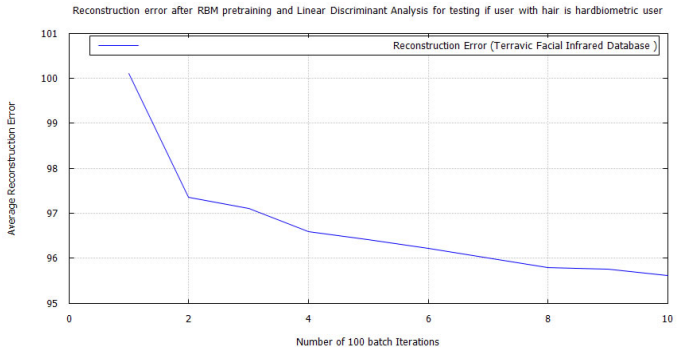


Figure 55: Error for facial hair algorithm testing hard biometric features for the presence of facial hair using RBM, LDA, and FWT for images from the Terravic Facial Infrared Database

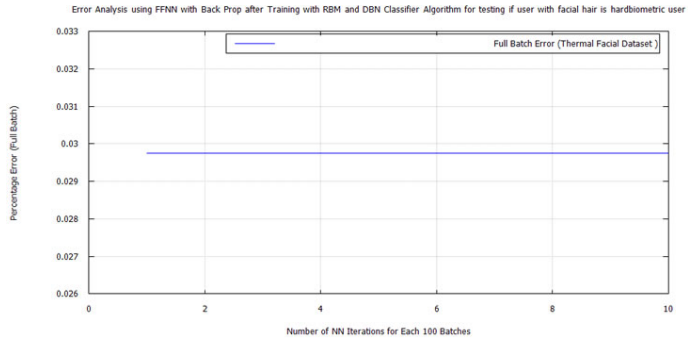


Figure 56: Error for facial hair algorithm testing if the hard biometric user is a user with facial hair using FFNN, DBN, RBM, LDA, and FWT for images from the Terravic Facial Infrared Database.

Table 8: FFNN Check for Facial Hair and Hard Biometric Traits

	Semantic Attribute	**FFNN error**
Training test	Hard-biometric face	0.02887
Training test	Faces with facial hair	0.02462
Verification test	Check if the user with hard-biometric features has facial hair as the main user does.	0.02975

7.2. Tests2 glass wearing and facial hair algorithms

The second test is conducted to verify that the user with facial hair, whose hard-biometric features have been confirmed, is putting on glasses. By using the X variables after LDA, of the glass wearing features for the user to be verified, and the y variables after LDA, of facial hair features. Based on facial hair attributes, the RBM, DBN, and FFNN tests were conducted to

verify that the user wearing glasses (presented to the sensor) is the authentic user. The FFNN result is shown, in comparison with those from the training test, in Table 9 below.

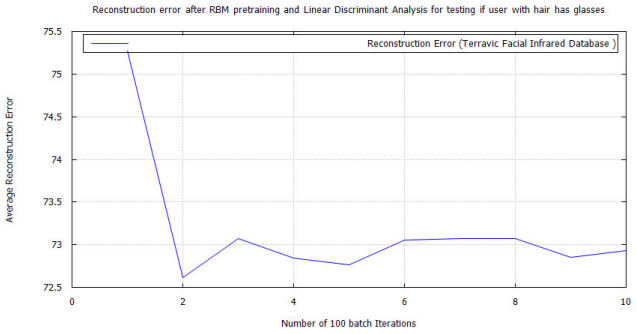


Figure 57: Error for facial hair algorithm testing glass wearing face for the presence of facial hair using RBM, LDA, and FWT for images from the Terravic Facial Infrared Database.

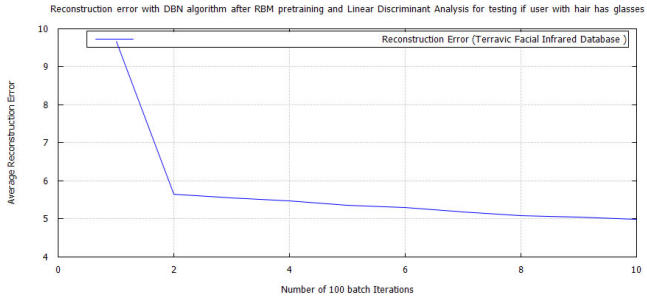


Figure 58: Error for facial hair algorithm testing glass wearing face for the presence of facial hair using DBN, RBM, LDA, and FWT for images from the Terravic Facial Infrared Database.

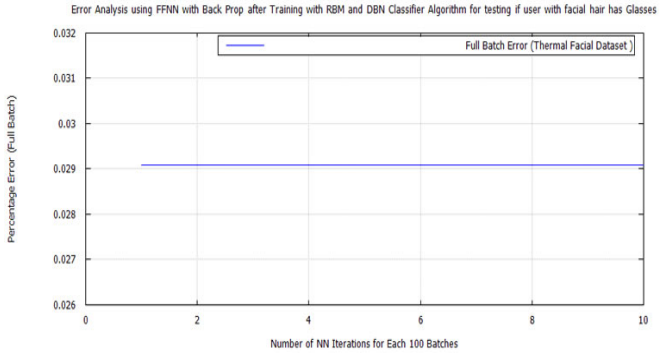


Figure 59: Error for facial hair algorithm testing glass wearing face for the presence of facial hair using FFNN, DBN, RBM, LDA, and FWT for images from the Terravic Facial Infrared Database

Table 9: FFNN Check for Facial Hair and Glass Wearing Face Traits

	Semantic Attribute	**FFNN error**
Training test	Faces with facial hair	0.02462
Training test	Faces in glasses	0.038695
Verification Test	Check if the user with facial hair has glasses	0.02908

7.3. Test 3, male gender and facial hair algorithms

The third test is conducted to verify that the user, with facial hair whose hard-biometric and glass wearing features have been confirmed, is masculine. By using the X variables after LDA, of the male gender features (from the sensor) of the user to be verified, and the Y variables after LDA of the facial hair features. RBM, DBN and FFNN tests are conducted to verify that the male, presented to the sensor, is the authentic user. The FFNN result is shown, in comparison with those from the training test, in Table 10.

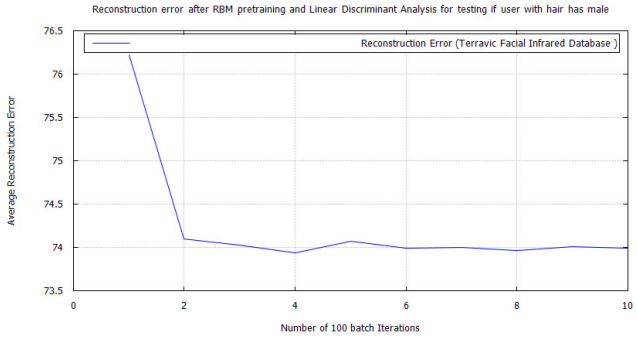


Figure 60: Error for facial hair algorithm testing male faces for the presence of facial hair using RBM, LDA, and FWT for images from the Terravic Facial Infrared Database

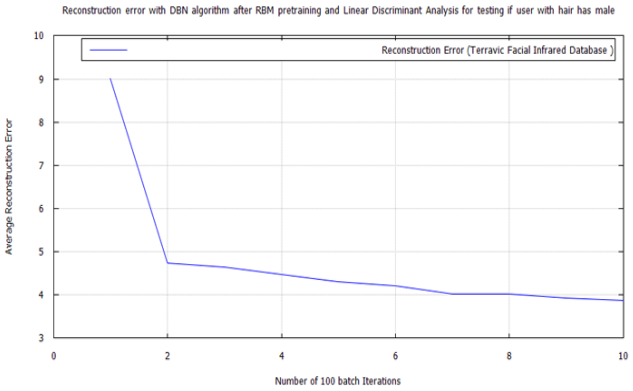


Figure 61: Error for facial hair algorithm testing male face for the presence of facial hair using DBN, RBM, LDA, and FWT for images from the Terravic Facial Infrared Database

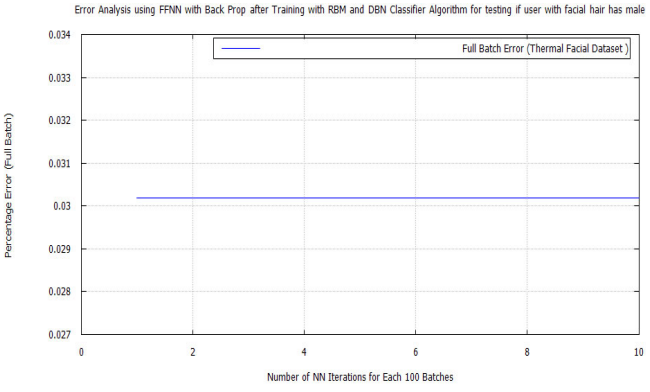


Figure 62: Error for male gender algorithm testing facial hair face for male gender using FFNN, DBN, RBM, LDA, and FWT for images from the Terravic Facial Infrared Database

Table 10: FFNN Check for Facial Hair and Male Gender Soft-Biometric Traits

	Semantic Attribute	**FFNN error**
Training test	Faces with facial hair	0.02462
Training test	Male faces	0.03
Verification Test	Check if a user with facial hair is male	0.03018

To ascertain that the person requesting verification is the authentic person. The results are summarized below.

7.4. Explanation

By comparing the FFNN error, from each of the cross-feature tests, to those from the algorithm. The trait check is seen to pass for each characteristic, based on comparison to the algorithm's test results. This is explained for each test as shown below.

7.4.1. Tests1, hard-biometric and facial hair algorithms

The similarity of the FFNN verification test results to those of the hard-biometric features, from the training test, show that they belong to the original user. This also shows that the hard-biometric trait, used in training the algorithm, has a close similarity to those possessed by the user being verified.

7.4.2. Tests2, glass wearing, and facial hair algorithms

The similarity of the FFNN verification test results to those of the face in glasses biometric features, from the training test, show that they belong to the original user. Since they have a value that is between that outputted by the glass algorithm and the one belonging to facial hair, as these two *soft-biometric* traits are present, in the X and y test data. Showing that the biometric trait, used in training the algorithm, has a close similarity to those possessed by the user being verified.

7.4.3. Test 3, male gender and facial hair algorithms

The similarity of the FFNN verification test result to those of the male gender biometric features, from the training test, show that they belong to the original user. Since they have a value that is closer to that of the male gender algorithm than the one belonging to facial hair. This also shows that the male gender biometric trait, used in training the algorithm, has a close similarity to those possessed by the user being verified.

7.4.4 Summary

From the tests, the first assumption was that the person being verified has facial hair and this was used to verify the other three traits, one after the other. The approach adds redundancy into the system with each trait used, to verify the other, one after the other. Such that if facial hair is absent, glasses would be and if glasses and facial hair are absent, the masculine features would be. Thus increasing robustness and system performance. In *identification mode*, for a large population size of millions, the technique can help to achieve high distinctiveness, high persistence, better performance, high acceptability and low potential for spoofing.

Conclusion

Multi-biometric person recognition, using thermal images to improve biometric security, is established, with images from Carl's database and the Terravic Facial Infrared Database. This was done using FWT, LDA and RBM pre-trained DBN with FFNN classifiers. The hard biometric attribute, of shape of head, was combined with outward attributes, such as face in glasses, face in a cap/head gear, face with facial hairs, plain face, female face and male face, using persons in the thermal facial datasets. These were learned, used for multi-biometric classification and then multi-biometric person recognition. Using the parallel mode of operation, through score fusion, a test user possessing these traits was verified with the technique. Showing how biometric system security can be fortified through a combination of hard and soft-biometric traits.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

References

[1] Moreno-Moreno, M., Fierrez, J., and Ortega-Garcia, J., "Millimeter- and Submillimeter-Wave Imaging Technologies for Biometric Purposes," in Proceedings of XXIV Simposium Nacional de Union Cientica Internacional de Radio, URSI, Spain, September 2009.

[2] D.L. Narayanan, R.N. Saladi, J.L. Fox, "Ultraviolet radiation and skin cancer," Int. J. Dermatol., vol. 49, no. 9, pp. 978–986, Sep. 2010.

[3] Moreno-Moreno, M., Fierrez, J., and Ortega-Garcia, J., "Biometrics Beyond the Visible Spectrum: Imaging Technologies and Applications," in Proceedings of BioID-Multicomm, LNCS 5707, 154–161, Springer, September 2009.

[4] E. Hurwitz, A. N. Hasan, and C. Orji, "*Soft biometric* thermal face recognition using FWT and LDA feature extraction method with RBM DBN and FFNN classifier algorithms," ICIIP, 2017, pp. 1 – 6, doi: 10.1109/ICIIP.2017.8313796.

[5] A.Meraoumia, S.Chitroub, A. Bouridane "Multimodal Biometric Person Recognition System based on Fingerprint & Finger-Knuckle-Print Using Correlation Filter Classifier", Communications (ICC), 2012 IEEE International Conference, Ottawa, Canada pages 820 - 824.

[6] Galbally, J., Marcel, S., & Fierrez, J. (2014). Biometric Antispoofing Methods: A Survey in Face Recognition. IEEE Access, 2, 1530-1552

[7] X. Geng, L. Wang, M. Li, Q. Wu, and K. Smith-Miles. Adaptive fusion of gait and face for human identification in video. In Applications of Computer Vision, 2008. WACV 2008. IEEE Workshop on, pages 1-6, Jan. 2008.

[8] L.Hong, A.k.Jain, S. Pankanti, "Can multibiometrics improve performance ?", Proc. AutoID'99, pp. 59-64, 1999-Oct

[9] IEEE OTCBVS WS Series Bench; Roland Miezianko, Terravic Research Infrared Database.

[10] V. Espinosa-Dur ´o, M. Faundez-Zanuy, and J. Mekyska, "A new face database simultaneously acquired in visible, near-infrared and thermal spectrums," Cognitive Computation, vol. 5, no. 1, pp. 119–135, 2013

[11] A. K. Jain, P. J. Flynn, and A. Ross. Handbook of Biometrics. Springer, 2007.

[12] H.T.F. Rhodes. Alphonse Bertillon: Father of scientific detection. Pattern Recognition Letters, 1956.

[13] L. Wen and G-D. Guo. A computational approach to body mass index prediction from face images. Image and Vision Computing, 31(5):392–400, 2013.

[14] R. Zewail, A. Elsafi, M. Saeb, and N. Hamdy. Soft and *hard biometrics* fusion for improved identity verification. In Proc. of IEEE International Midwest Symposium on Circuits and Systems, volume 1, pages I – 225–8, 2004.

[15] D. Reid and M. Nixon. Imputing human descriptions in semantic biometrics. In Proc. of ACM Workshop on Multimedia in Forensics, Security and Intelligence, 2010.

[16] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," IEEE Trans. Circuits Syst. Video Technology, Special Issue Image- and Video-Based Biomet., vol. 14, no. 1, pp. 4–20, Jan. 2004.

[17] R.M. Bolle, S. Pankanti, and N.K. Ratha, "Evaluation Techniques for Biometrics-Based Authentication Systems (FRR)," Proc. 15th Int'l Conf. Pattern Recognition, vol. 2, pp. 831-837, Sept. 2000.

[18] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar. Attribute and simile classifiers for face verification. In Proc. of International Conference on Computer Vision, 2009.

[19] S. Z. Li and A. K. Jain, Eds., Handbook of Face Recognition. New York: Springer Verlag, 2004.

[20] D. Reid and M. Nixon. Using comparative human descriptions for soft biometrics. In Proc. of International Joint Conference on Biometrics, 2011.

[21] S. Samangooei, B. Guo, and Mark S. Nixon. The use of semantic human description as a soft biometric. In Proc. of IEEE International Conference on Biometrics: Theory, Applications and Systems, 2008.

[22] A.K. Jain, S.C. Dass, and K. Nandakumar. *Soft biometric* traits for personal recognition systems. In Proceedings of ICBA, pages 1–40. Springer, 2004.

[23] A. Dantcheva, C. Velardo, A. D'Angelo, and J.-L. Dugelay. Bag of *soft biometrics* for person identification. New trends and challenges. Multimedia Tools and Applications, 51:739–777, 2011.

[24] D. Bhattacharjee, A. Seal, S. Ganguly, M. Nasipuri, and D.K. Basu, "A Comparative Study of Human Thermal Face Recognition Based on Haar Wavelet Transform and Local Binary Pattern," Computational Intelligence and Neuroscience, vol. 2012, Article ID 261089, 12 pages, 2012. oi:10.1155/2012/261089

[25] L. Hong, A. K. Jain, S. Pankanti, "Can multibiometrics improve performance ?", Proc. AutoID'99, pp. 59-64, 1999-Oct

[26] B. Martinez, X. Binefa, and M. Pantic. Facial component detection in thermal imagery. In Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops, pages 48–54, 2010

[27] C. Herrmann, T. Müller, D. Willersinn, J. Beyerer, "Real-time person detection in low-resolution thermal infrared imagery with MSER and CNNs", in Proc. Electro-Optical and Infrared Systems, SPIE, 2016

[28] A. Krizhevsky, I. Sutskever and G. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In Advances in Neural information Processing Systems 25, pages 1106-1114, 2012

[29] W. Shangfei, L. Zhilei, L. Siliang, L. Yanpeng, W. Guobing, P. Peng, C. Fei and W. Xufa, "A Natural Visible and Infrared Facial Expression Database for Expression Recognition and Emotion Inference", IEEE Transactions on Multimedia, vol.12, no.7, pp.682-691, Nov. 2010

[30] K. Tai, S. Blain, and T. Chau, "A review of emerging access technologies for individuals with severe motor impairments," Assist. Technol., vol. 20, pp. 204–219, 2008

[31] Bhowmik, M.K., Saha, K., Majumder, S., Majumder, G., Saha, A., Sarma, A.N., Bhattacharjee, D., Basu, D.K., Nasipuri, M., Corcoran, Peter M. 'Thermal infrared face recognition – a biometric identification technique for robust security system'. Reviews, Refinements and New Ideas in Face Recognition. 2011, 07, Zieglergase, Vienna, Austria, Europe: InTech Open Access Publisher, Vienna Office, ISBN: 978-953-307-368-2, (open access publisher of scientific books and journals)

[32] A. Seal, S. Ganguly, D. Bhattacharjee, M. Nasipuri, D. K. Basu, "Minutiae Based Thermal Human Face Recognition using Label Connected Component Algorithm," Procedia Technology, In 2nd International Conference on Computer, Communication, Control and Information Technology (C3IT-2012), February 25 - 26, 2012. multilayer perceptron

[33] S. Cho, L. Wang, and W. J. Ong. Thermal imprint feature analysis for face recognition. ISIE, pages 1875–1880, 2009.

[34] O. Arandjelovi´c, R. Hammoud, and R. Cipolla. On person authentication by fusing visual and thermal face biometrics. In Proc. IEEE Conference on Advanced Video and Singal Based Surveillance, pages 50–56, 2006.

[35] M. Piccardi, "Background subtraction techniques: a review," in Proc. of the IEEE International Conference on Systems, Man and Cybernet- ics (SMC), vol. 4, pp. 3099-3104, IEEE, Oct. 2004.

[36] S. Se and M. Brady, "Ground plane estimation, error analysis and applications," Robotics and Autonomous Systems, vol. 39, no. 2, pp. 59-71, 2002.

[37] K. C. Sofiane Yous, Hamid Laga, "People detection and tracking with world-z map from a single stereo camera," in Proc. of the International Workshop on Visual Surveillance (VS) - (ECCV), (Marseille, France), October 2008.

[38] A. Broggi, M. Bertozzi, A. Fascioli, and M. Sechi, "Shape-based pedes-trian detection," in Proc. of the IEEE Intelligent Vehicles Symposium(IV), pp. 215-220, 2000.

[39] D. Vaquero, R. Feris, D. Tran, L. Brown, A. Hampapur, and M. Turk. Attribute-based people search in surveillance environments. In Proc. of IEEE Workshop on Applications of Computer Vision, 2009.

[40] B. Wu and R. Nevatia, "Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part de-tectors," in Proc. of the International Conference on Computer Vi-sion(ICCV), 2005.

[41] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms I. Communications on Pure and Applied Mathematics, 44 (2):141–183, March 1991

[42] M. Jones, P. Viola, P. Viola, M. J. Jones, D. Snow, and D. Snow, "De-tecting pedestrians using patterns of motion and appearance," in Proc. of the International Conference on Computer Vision(ICCV), pp. 734-741, 2003.

[43] Q. Zhu, M.-C. Yeh, K.-T. Cheng, and S. Avidan, "Fast human de-tection using a cascade of histograms of oriented gradients," in Proc. of the IEEE Conference on Computer Vision and Pattern Recogni-tion(CVPR), (Washington, DC, USA), pp. 1491-1498, IEEE Computer Society, 2006.

[44] A. Mohan, C. Papageorgiou, and T. Poggio, "Example based object detection in images by components," IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 23, pp. 349-361, 2001.

[45] B. Wu and R. Nevatia, "Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet based part de-tectors," International Journal of Computer Vision(IJCV), vol. 75, pp. 247-266, Nov. 2007

[46] L.Trujillo, G. Olague, R.Hammoud and B. Hernandez. Automatic Feature Localization in Thermal Images for Facial Expression Recognition. Computer Vision and Pattern Recognition-Workshops, IEEE Computer Society Conference, p. 14, 2005.

[47] Martinez B, Binefa X, Pantic M. Facial Component Detection in Thermal Imagery. Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE Computer Society Conference, p. 48-54, 2010.

[48] F.Q Al-Khalidi, R. Saatchi, D. Burke, H. Elphick. Tracking human face features in thermal images for respiration monitoring. Int Conf Computer Systems and Applications(AICCSA), p. 1-6, 2010.

[49] S. Prabhakar, S. Pankanti, and A. K. Jain, "Biometric recognition: Security and privacy concerns," IEEE Security Privacy Mag., vol. 1, no. 2, pp. 33–42, 2003

[50] R.M. Bolle, S. Pankanti, and N.K. Ratha, "Evaluation Techniques for Biometrics-Based Authentication Systems (FRR)," Proc. 15th Int'l Conf. Pattern Recognition, vol. 2, pp. 831-837, Sept. 2000.

[51] Joshua C. Klontz and Anil K. Jain. A case study on unconstrained facial recognition using the boston marathon bombings suspects. Technical Report, (MSU-CSE-13-4), 2013

[52] S. Z. Li and A. K. Jain, Eds., Handbook of Face Recognition. New York: Springer Verlag, 2004.

[53] International Biometric Group, "Independent Testing of Iris RecognitionTechnology May 2005 [Online]. Available: http://www.biometricgroup. com/reports/public/reports/ITIRT_report.htm.

[54] A.K. Jain, S. C. Dass, and K. Nandakumar. Can *soft biometric* traits assist user recognition? In Proceedings of SPIE, volume 5404, pages 561–572, 2004

[55] W. J. Scheirer, N. Kumar, K. Ricanek, P. N. Belhumeur, and T. E. Boult. Fusing with context: a Bayesian approach to combining descriptive attributes. In Proc. of International Joint Conference on Biometrics, 2011

[56] S. Denman, C. Fookes, A. Bialkowski, and S. Sridharan. Softbiometrics: Unconstrained authentication in a surveillance environment. In Proc. of International Conference on Digital Image Computing: Techniques and Applications, pages 196–203, 2009

[57] N. Holighaus, Z. Pr˚u˘sa, C. Wiesmeyr, "Designing tight filter bank frames for nonlinear frequency scales, sampling Theory and Applications," 2015. [online]. Available: http://ltfat.github.io/notes/ltfatnote039.pdf.

[58] P. Wagner, "Bytefish face recognition algorithms for MATLAB/GNU Octave and Python," 2015. [Online]. Available: https://github.com/bytefish/facerec.

[59] R. B. Palm, "Prediction as a candidate for learning deep hierarchical models of data," 2012. [Online]. Available: http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=6284

[60] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms I. Communications on Pure and Applied Mathematics, 44(2):141–183, March 1991.

[61] L. Sirovitch and M. Kirby, "Low-Dimensional Procedure for the Characterization of Human Faces," J. Optical Soc. of Am. A, vol. 2, pp. 519-524, 1987.

[62] P.L. Søndergaard, B. Torrésani, P.Balazs. The Linear Time-Frequency Analysis Toolbox. International Journal of Wavelets, Multiresolution Analysis and Information Processing, 10(4), 2012.

[63] M. Turk and A. Pentland, "Eigenfaces for Recognition," J. Cognitive Neuroscience, vol. 3, no. 1, 1991.

[64] R.A. Fisher, "The Use of Multiple Measures in Taxonomic Problems," Ann. Eugenics, vol. 7, pp. 179-188, 1936.

[65] V. Nair and G. E. Hinton. Rectified linear units improve restricted boltzmann machines. In Proc. 27th International Conference on Machine Learning, 2010.

[66] G. E. Hinton, S. Osindero, and Y. Teh, "A fast learning algorithm for deep belief nets," Neural computation, vol. 18, no. 7, pp. 1527–1554, 2006.

[67] A. K. Jain, A. Ross, and S. Pankanti, "Biometrics: a tool for information security," IEEE Transactions on Information Forensics and Security, vol. 1, no. 2, pp. 125–143, 2006.

[68] M. Madry-Pronobis. Automatic gender recognition based on audiovisual cues. Master Thesis, 2009.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 74-79 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

In Body Antenna for Monitoring and Controlling Pacemaker

Sourav Sinha*, Ta-Seen Reaz Niloy, Raja Rashidul Hasan, Md. Abdur Rahman

Department of Electrical & Electronic Engineering, American International University-Bangladesh (AIUB), Dhaka-1229, Bangladesh

ARTICLE INFO

Article history:
Received: 13 December, 2019
Accepted: 05 February, 2020
Online: 09 March, 2020

Keywords:
Pacemaker
Wireless monitoring
Micro-strip antenna
Perfect Electric Conductor
Specific Absorption Rate
2/3 muscle-equivalent phantom
Wideband
Directivity
Rogers

ABSTRACT

This paper is an extension of work originally presented in 2019 International Conference on Automation, Computational and Technology Management (ICACTM). A micro-strip patch in-body designed antenna is constructed on pacemaker to monitor and control the pacemaker wirelessly. The antenna is intended for ISM (Industrial, Scientific, and Medical) band (2.4 GHz to 2.48 GHz). A perfect electric conductor (PEC) is considered as pacemaker body and used as the ground of the propounded antenna having dimensions 40 x 30 x 10 mm^3. The patch material is chosen Copper having dimensions 35 x 22 x 0.1 mm^3 and covered up with substrate material Rogers R03010 (loss tangent δ = 0.0035 and dielectric constant, ε_r = 10.2) with thickness of 1.55 mm to make it compatible in human body. The designed antenna is placed and analyzed in 2/3 muscle equivalent phantom by changing the depth of the antenna. Results disclose that operating frequency is 2.464 GHz with reflection coefficient -28.37 dB. The antenna maintains frequency range from 1.8075 GHz to 3.445 GHz, which represents wide bandwidth of 1.6375 GHz. To ensure the human body safety, specific absorption rate is analyzed and found 0.937 W/Kg for 10g tissue at operating frequency, which makes it biocompatible. The surface current distribution, Voltage Standing Wave Ratio, Current density, far-field radiation characteristics, radiation efficiency, and total efficiency are investigated to analyze the effect and performance of the designed antenna. CST Microwave Studio is used for simulation and analysis the parameters of the antenna.

1. Introduction

Nowadays in our modern scenario, heart attack is very familiar disease caused by loss of blood supply. It appears when the heart is beating too rapidly or when the body is not receiving enough blood [1]. A small electrical medical device called artificial implantable pacemaker provides an electrical pulse to the heart with proper intensity by the electrodes [2]. The main objective of pacemaker is to regulate irregular heartbeats (known as arrhythmias) at normal rate. Many organizations come up with distinctive features, such as wireless communication and monitoring assistance [3]. This monitoring and controlling system is extremely relevant because it decreases the amount of eye to eye visits with doctors, as well as the physical and mental stress of patients. Since it can communicate wirelessly, there is no need to cut the skin, which can prevent infection from a germ in a medical diagnosis, but the antenna is needed to transmit the data wirelessly.

A micro-strip patch antenna's geometric shape contains a shedding layer on first part of the dielectric substrate with a ground plane on the next. Radiating element can be square, circular, semicircular, triangular etc. [4]. In present days, a number of antenna design have been proposed on several researches for the different applications. With many advantages, this antenna can communicate with others wireless system spontaneously [5, 6]. Easy to manufacture, lightweight, low profile, low cost, replaceable and high efficiency is such advances. The popularity of patch antenna in microwave communication is raised nowadays that requires semispherical coverage. Safety of a patient with a compact size, as well as wider bandwidth and radiation efficiency is the main concern to design an antenna [7, 8]. Various types of antenna have been developed, e.g. meander line, monopole, and loop antenna for wireless monitoring [9, 10].

ISM or Industrial, Scientific and Medical band (2.4 GHz - 2.4835 GHz) and MICS or Medical Implant Communication System band (402 MHz to 405 MHz) specified for bio-suitable antenna by U.S. Federal Communications Commission and European Radio communications Committee (ERC) in wireless medical telemetry service (WMITS) [11]. ISM band has shorter wavelength due to its high frequency area range apart from MICS

*Sourav Sinha, +8801711190232, souravsinha272@gmail.com

band has wider wavelength due to its smaller frequency. Therefore, ISM band is extra reasonable than MICS to design an antenna on pacemaker [12].

In this article, a design of bio-implantable micro-strip patch antenna is propounded for monitoring and controlling pacemaker, which operates at ISM band. The pacemaker with the antenna placed inside 2/3 muscle equivalent phantom, then the characteristics are analyzed. The performance is also analyzed by changing depth of the pacemaker in 2/3 muscle equivalent phantom. The pacemaker box considered as box of PEC (perfect electric conductor) as well as ground of the proposed antenna [13]. Rogers R03010 ($\varepsilon_r = 10.2$) covers the antenna from all the side can also called superstrate and substrate of the antenna for its flexibility and durability characteristics [14]. All the characteristics of the propounded antenna observed in CST microwave studio and discussed in this article.

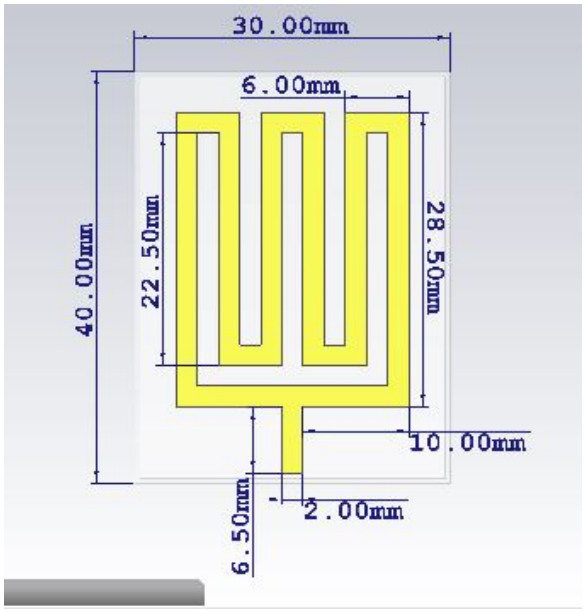


Figure 1: Propounded micro-strip patch antenna front view with dimension.

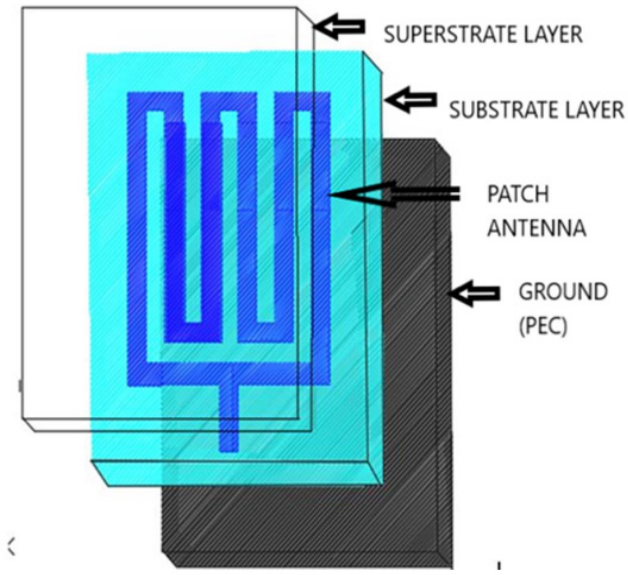


Figure 2: Antenna Structure

2. Structure and Design Method

2.1. Design of Antenna

The propound micro-strip patch antenna has the length 35 mm and width 22 mm respectively, which overcome the most challenging issues - biocompatibility and size. The designed antenna is consisting superstrate, patch, substrate, and ground. Copper is chosen for patch with 0.1 mm of thickness. The pacemaker box is built with PEC or perfect electric conductor with thickness of 0.1 mm from all sides and it acts as ground of the antenna. The substrate and superstrate are made with Rogers R03010 that covers full patch material with 1.55 mm thickness.

The geometrical front view of antenna shows all the dimensions of patch, which are placed in Figure 1. The patch is located in the middle of the superstrate and substrate with 35 x 22 x 0.1 mm^3 to reduce the effects of a high conductive human tissue as well as to avoid shortening the antenna [12-15]. Having 6.5 mm length and 2 mm width, feed line is also shown. All the dimensions are organized in Table 1.

Table 1: The Antenna Parameters

Antenna Part	Material	Parameter	Value (mm)
Superstrate	Rogers	Length	40
		Width	30
		Thickness	1.55
Patch	Copper	Length	35
		Width	22
		Thickness	0.1
Substrate	Rogers	Length	40
		Width	30
		Thickness	1.55
Ground	PEC	Length	40
		Width	30
		Thickness	0.1

2.2. Design of Pacemaker Casing

Figure 3 shows casing of Pacemaker and the location of the antenna with waveguide port. The antenna patch covers with substrate and superstrate is placed on the top surface of the antenna. The waveguide port is connected underneath the feed line from where the power of input is given. The red part states the port of the waveguide as shown in Figure 3. Dimensions of the pacemaker case are tabulated in Table 2.

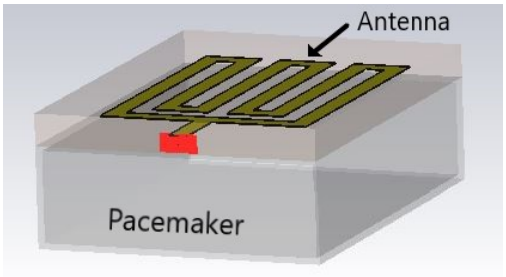


Figure 3: Antenna with Pacemaker Case Showing Waveguide Port

Table 2: Pacemaker Case Dimension

Name of Parameters	Value (mm)
Pacemaker Length	30
Pacemaker Width	40
Pacemaker Thickness	10
Pacemaker Case Thickness	0.1

2.3. Structure of Body Phantom

As shown in Figure 4, the pacemaker with surfaced antenna is placed inside the 2/3 muscle-equivalent phantom. The interval between antenna and the phantom's surface is denoted with "d" and it varies depending on gender and age.

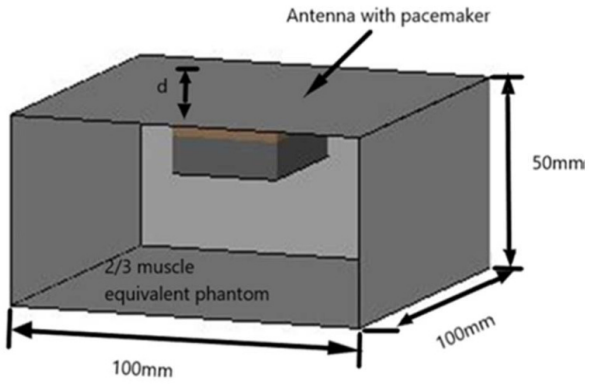


Figure 4: 2/3 muscle-equivalent phantom model pacemaker inside

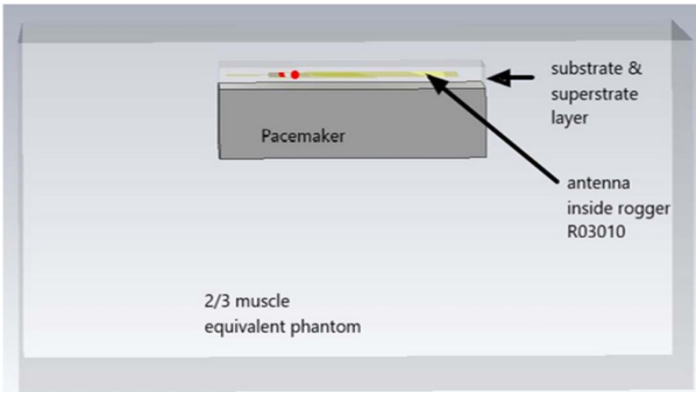


Figure 5: Cross sectional view of 2/3 muscle phantom with pacemaker inside.

Cross sectional view of the designed antenna with pacemaker inside the 2/3 muscle equivalent phantom is shown in Figure 5. The positioning pattern of pacemaker is clearly described.

3. Characteristics Analysis of Designed Antenna

3.1. Reflection coefficient or S_{11} parameter

Reflection coefficient or S_{11} parameter measures the quantity of power that radiated or reflected from an antenna [16]. The return loss is noticed after implanting the proposed antenna inside 2/3 muscle equivalent phantom model at resonant frequency. In Figure 6, the X-axis is in GHz range, which represents the frequency and the Y-axis is in dB scale, which constitutes the return loss. The resonant frequency or operating frequency of the propounded antenna is found 2.464 GHz, which is in ISM band and that turns the antenna into biocompatible and also return loss of -28.369 dB that indicates better performance by maximum radiation of the antenna [14]. The bandwidth of the propounded antenna is observed 1.6375 GHz (1.8075 GHz to 3.445 GHz) by sketching a linear line in -10 dB, which is enough sufficient to implant on human body [14].

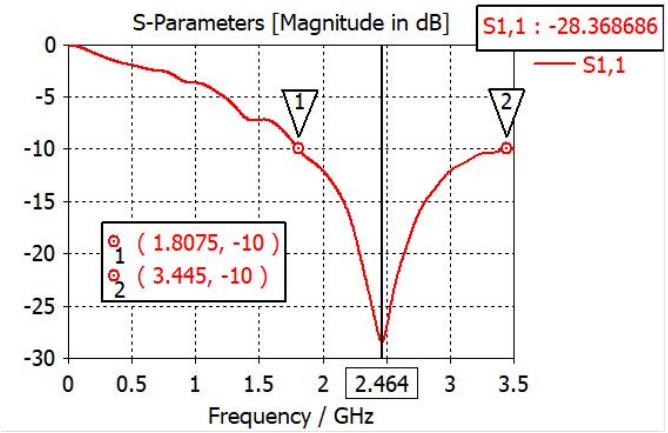


Figure 6: Return loss (S_{11} parameter) of the antenna inside 2/3 muscle equivalent phantom model.

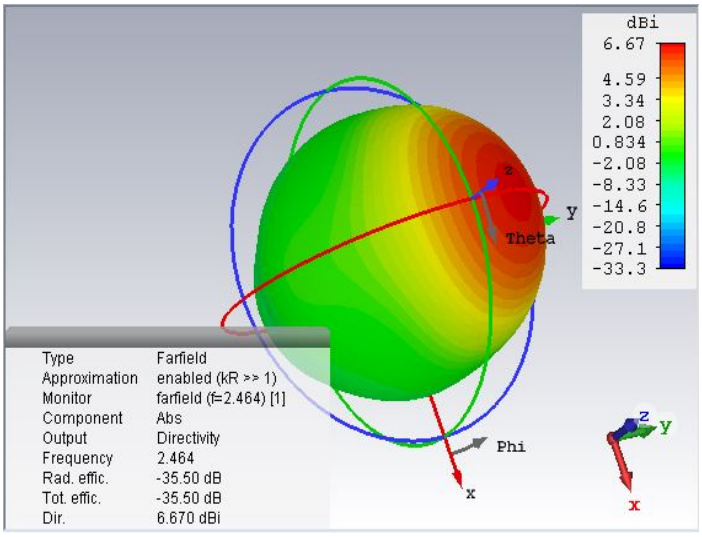
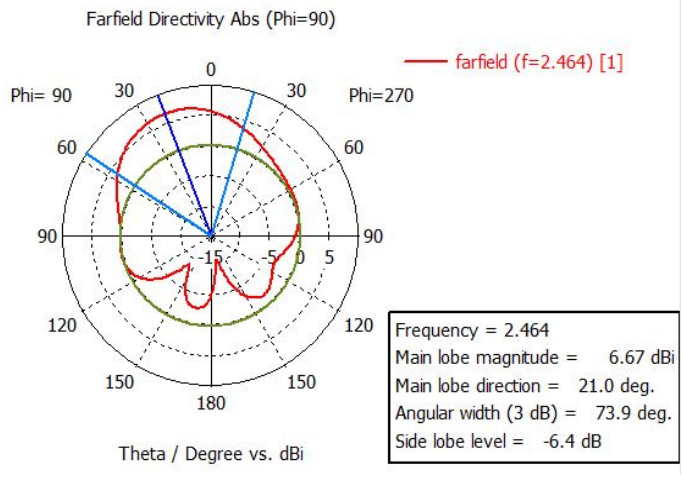


Figure 7: Far-field radiation pattern view (3D) of the propounded antenna inside 2/3 muscle equivalent phantom model.



3.2. Far-field radiation pattern

The radiation characteristics of the proposed antenna is explained in Figure 7. As the antenna is designed for monitoring and controlling pacemaker so unidirectional directivity is maintained to get the best values because pacemaker is a static device inside human body. Various parameters such as the directivity, total efficiency and the radiation efficiency is got 6.67 dBi, -35.5 dB and -35.5 dB respectively at the resonance frequency of 2.464 GHz and represented the radiation characteristics of the designed antenna. Figure 8 is displayed the far-field radiation pattern's polar view of the proposed antenna with main lobe magnitude 6.67 dBi.

3.3. Voltage standing wave ratio or VSWR

The voltage standing wave ratio is a function of the coefficient of reflection and of the intensity reflected from antenna [16]. For better performance, VSWR should be in between 1 to 2. In Figure 9, the X-axis is in GHz range, which represents the frequency and the Y-axis is ratio, which represents VSWR. In addition, it founds 1.08 at resonance frequency 2.464 GHz.

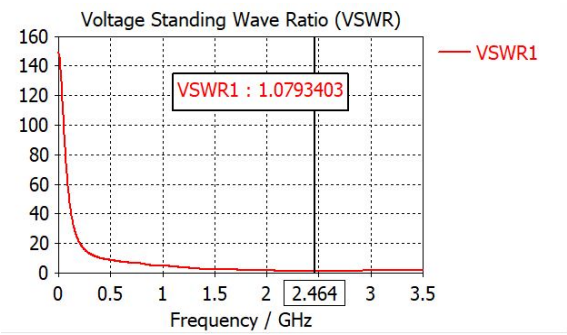


Figure 9: Voltage standing wave ratio (VSWR) of the antenna inside 2/3 muscle equivalent phantom model.

3.4. Specific Absorption Rate (SAR)

SAR is measured by the radiation of the surrounding tissue [17] that is used for safety purpose. According to FCC, in 10g tissue and 1mW input power, SAR must be less than 2 W/Kg to ensure higher safety [18, 19]. From Figure 10, Maximum Specific Absorption Rate (SAR) for the proposed antenna is noticed 0.937 W/kg at operating frequency for 10g tissue.

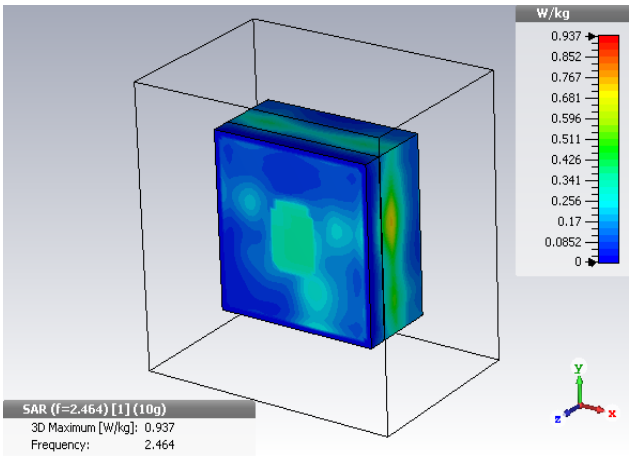


Figure 10: SAR distribution of the propounded antenna for 10g tissue in 1mW of input power.

3.5. Current density

Figure 11 shows the current density (abs) of the antenna at resonance 2.464 GHz. In addition, the current magnitude is not same in every position of the antenna patch. The current density in patch lies between 1017 A/m^2 to 1176 A/m^2. In addition, the maximum current density is found 5851 A/m^2.

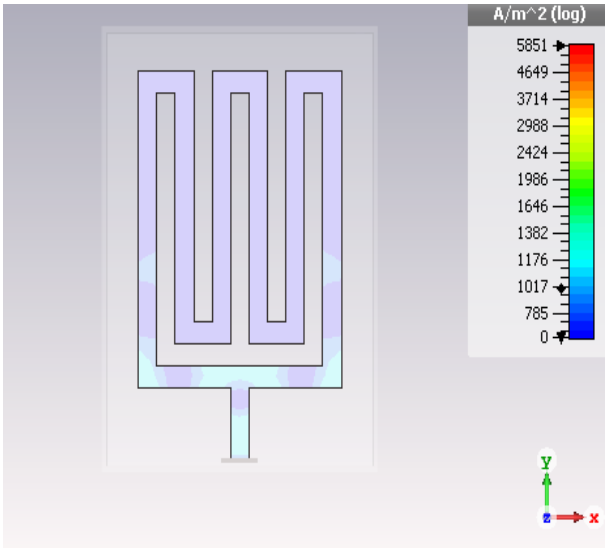


Figure 11: Current density of the designed antenna at 2.464 GHz.

3.6. Surface current distribution

The surface current distribution of the designed antenna at resonance frequency 2.464GHz is shown on Figure 12. The maximum current is found 52.5 A/m in design while the peck current on patch is found about 14.3 A/m. Figure 12 is also clearly described that the current is high on the lower edges which close to feeding point and low on the upper edges.

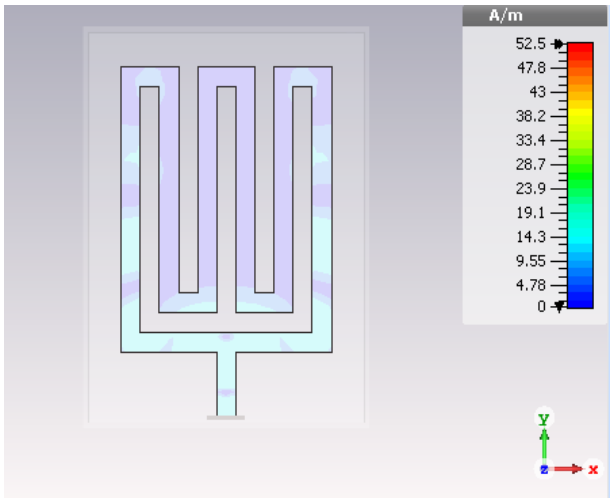


Figure 12: Surface current distribution of the designed antenna at 2.464 GHz.

4. Comparison Analysis

4.1. For various depth

Depending on the gender and age, the depth of the pacemaker is varied and analyzed. The depth defines the distance between the phantom surface and the superstrate surface. As shown in Figure

13, the reflection coefficient or S_{11} and the operating frequency is slightly shifted with the change of depth (d). Table 3 is shown the total changelog for different depth of pacemaker.

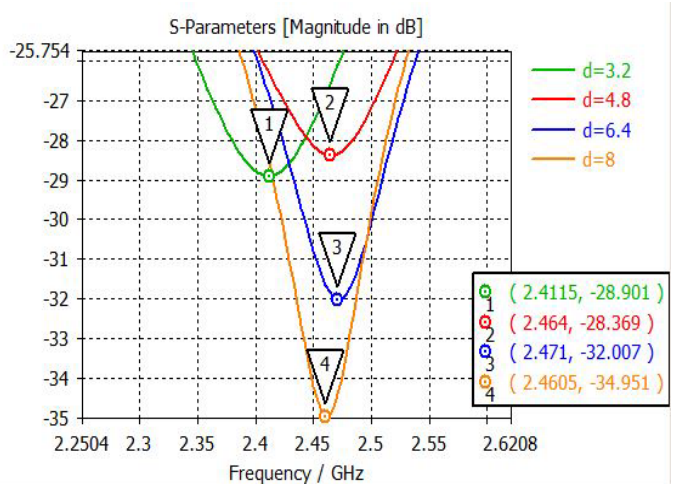


Figure 13: Return loss or S_{11} parameter of the antenna for various depth

Table 3: Changelog with different depth

Position	Depth (mm)	Operating Frequency	S_{11}
1	3.2	2.4115	-28.901
2	4.8	2.464	-28.369
3	6.4	2.471	-32.007
4	8	2.4605	-34.951

4.2. By changing material

The designed antenna parameters are also analyzed with two patch materials Copper, Gold and two substrate-superstrate materials Roger and FR-4 variations. As shown in Figure 14, the resonance is found 2.464 GHz frequency for all possible combinations and return loss or S_{11} parameters are almost same in the variations. Table 4 displays all the parameters findings by changing the materials.

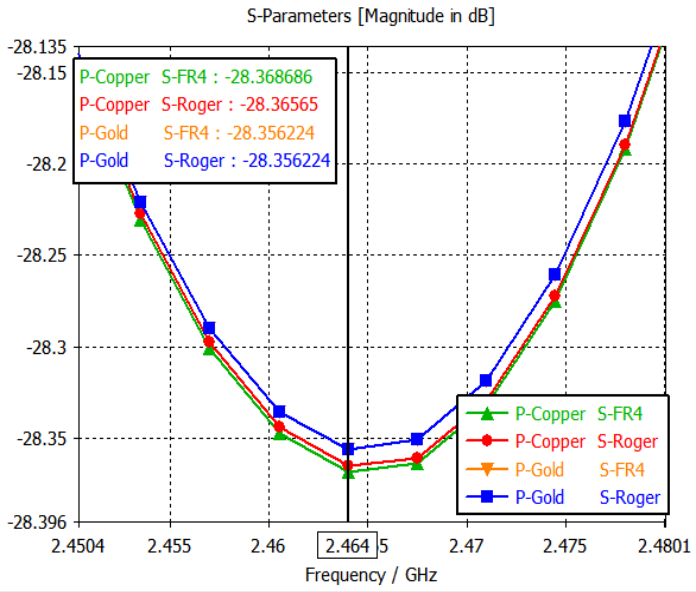


Figure 14: Return loss or S_{11} parameter of the antenna for different material.

Table 4: Changelog with different materials

Patch	Substrate	S_{11}	Directivity	VSWR	SAR
Copper	FR-4	-28.37	6.67 dBi	1.07934	0.937
Copper	Roger	-28.36	6.67 dBi	1.07937	0.937
Gold	FR-4	-28.35	6.67 dBi	1.07946	0.937
Gold	Roger	-28.35	6.67 dBi	1.07946	0.937

5. Conclusion

In this article, a micro-strip patch antenna is proposed for monitoring and controlling pacemaker. Biocompatibility have been checked by embedded the propounded antenna with pacemaker inside 2/3 muscle equivalent phantom. The depth is varied and analyzed to ensure same characteristics for any gender and any age's people. S_{11} or reflection coefficient is found -28.37 and the resonance frequency is found 2.464 GHz for 4.8 mm depth. Pacemaker is one of the vital device nowadays for heart patients, especially for arrhythmias patients. Any kind of interruption in working of pacemaker can turn a patient into death. Therefore, monitoring and controlling pacemaker is become most important issue for pacemaker holders. In conclusion, according to the antenna performance, the propounded antenna is biocompatible and applicable for medical application with proper output.

References

[1] R. R. Hasan, M. A. Rahman, S. Sinha, M. N. Uddin, T. R. Niloy, "In Body Antenna for Monitoring Pacemaker" in 2019 International Conference on Automation, Computational and Technology Management (ICACTM), London, United Kingdom, 2019. doi: 10.1109/ICACTM.2019.8776836

[2] Y. Hao, Y. Li, D. Liao, L. Yang, "Seven times replacement of permanent cardiac pacemaker in 33 years to maintain adequate heart rate: a case report" Annals of translational medicine, 3(21), 2015. doi: 10.3978/j.issn.2305-5839.2015.11.35

[3] P. Soontornpipit, C. M. Furse, Y. C. Chung, "Miniaturized biocompatible microstrip antenna using genetic algorithm," IEEE Transactions on Antennas and Propagation, 53(6), 1939-1945, 2005. doi: 10.1109/TAP.2005.848461

[4] H. Werfelli, K. Tayari, M. Chaoui, M. Lahiani, H. Ghariani, "Design of rectangular microstrip patch antenna" in 2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Monastir, 2016. doi: 10.1109/ATSIP.2016.7523197

[5] C. Liu, Y. Guo, H. Sun, S. Xiao, "Design and Safety Considerations of an Implantable Rectenna for Far-Field Wireless Power Transfer" IEEE Transactions on Antennas and Propagation, 62(11), 5798-5806, 2014. doi: 10.1109/TAP.2014.2352363

[6] S. Gollakota, H. Hassanieh, B. Ransford, D. Katabi, K. Fu, "They can hear your heartbeats: non-invasive security for implantable medical devices" Proc. ACM SIGCOMM 2011, 41(4), 2-13, 2011. doi: 10.1145/2018436.2018438

[7] W. El Hajj, C. Person, J. Wiart, "A Novel Investigation of a Broadband Integrated Inverted-F Antenna Design; Application for Wearable Antenna" IEEE Transactions on Antennas and Propagation, 62(7), 3843-3846, 2014. doi: 10.1109/TAP.2014.2318061

[8] Z. Duan, Y. Guo, M. Je, D. Kwong, "Design and in Vitro Test of a Differentially Fed Dual-Band Implantable Antenna Operating at MICS and ISM Bands" IEEE Transactions on Antennas and Propagation, 62(5), 2430-2439, 2014. doi: 10.1109/TAP.2014.2309130

[9] J. Kim and Y. Rahmat-Samii, "Implanted antennas inside a human body: simulations, designs, and characterizations" IEEE Transactions on Microwave Theory and Techniques, 52(8), 1934-1943, 2004. doi: 10.1109/TMTT.2004.832018

[10] P. Soontornpipit, C. M. Furse, Y. C. Chung, "Design of implantable microstrip antenna for communication with medical implants" IEEE Transactions on Microwave Theory and Techniques, 52(8), 1944-1951, 2004. doi: 10.1109/TMTT.2004.831976

[11] K. S. Nikita, Handbook of Biomedical Telemetry, IEEE press, 2014.

[12] H. Usui, M. Takahashi, K. Ito, "Radiation characteristics of an implanted cavity slot antenna into the human body" in 2006 IEEE Antennas and

Propagation Society International Symposium, Albuquerque, NM, 2006. doi: 10.1109/APS.2006.1710726

[13] T. Houzen, M. Takahashi, K. Saito, K. Ito, "Implanted Planar Inverted F-Antenna for Cardiac Pacemaker System" in 2008 International Workshop on Antenna Technology: Small Antennas and Novel Metamaterials, Chiba, 2008. doi: 10.1109/IWAT.2008.4511351

[14] M. Islam, R.R. Hasan, M. Haque, S. Ahmad, K. Mazed, M. Islam, "In-Body Antenna for Wireless Capsule Endoscopy at MICS Band" in Computing Conference 2018, London, United Kingdom, 2018. doi: 10.1007/978-3-030-01177-2_59

[15] S. Sinha, R. R. Hasan, M. A. Rahman, M. T. Ali, M. N. Uddin, "Antenna Design for Biotelemetry System" in 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), Dhaka, Bangladesh, 2019. doi: 10.1109/ICREST.2019.8644307

[16] R. R. Hasan, M. A. H. Shanto, S. Howlader, S. Jahan, "A novel design and miniaturization of a scalp implantable circular patch antenna at ISM band for biomedical application" in 2017 Intelligent Systems Conference (IntelliSys), London, United Kingdom, 2017. doi: 10.1109/IntelliSys.2017.8324286

[17] N. A. Islam, F. Arifin, "Performance analysis of a miniaturized implantable PIFA antenna for WBAN at ISM band" in 2016 3rd International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), Dhaka, 2016. doi: 10.1109/CEEICT.2016.7873145

[18] E. Safety, S. C. Committee, N. Radiation, I. S. Board, "IEEE Recommended Practice for Measurements and Computations of Radio Frequency Electromagnetic Fields With Respect to Human Exposure to Such Fields 100 kHz –300 GHz" Measurement, 2002, 2008.

[19] J. H. Bernhardt, "The New ICNIRP Guidelines: Criteria, Restrictions, and Dosimetric Needs" Radio Frequency Radiation Dosimetry and Its Relationship to the Biological Effects of Electromagnetic Fields. NATO Science Series (Series 3: High Technology), 82, 2000. doi: 10.1007/978-94-011-4191-8_56

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 77-85 (2020)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Routing Protocols for VANETs: A Taxonomy, Evaluation and Analysis

Amina Bengag*, Asmae Bengag, Mohamed Elboukhari

MATSI lab, ESTO, Mohammed First University, Oujda, Morocco

ARTICLE INFO

Article history:
Received: 21 September, 2019
Accepted: 27 December, 2019
Online: 15 January, 2020

Keywords:
VANETs
Routing protocols
SUMO
NS-3
OSM
Urban scenario
Greedy forwarding
Angle direction

ABSTRACT

VANET as a subclass of MANET is composed of a set of vehicles equipped with wireless transceivers, to build dynamic networks without the need of any pre-existing infrastructure. Over the last few decades, the area of routing protocols in VANETs has been extensively studied. Nevertheless, this area remains even more challenging due to some features of VANETs, such as the high speed of vehicles, the often-disconnected links and the particular mobility pattern. Routing protocols in VANETs could be splitted into four categories: topology, position, multicast and broadcast-based routing protocols. In this paper, we provide a novel detailed taxonomy of routing protocols in VANETs then we present the advantages and drawbacks of each category. Moreover, we clear up the techniques adopted by each of the most popular routing protocols based on the vehicles' position and the topology of the networks. To explore the strengths and weaknesses of each routing protocols, basing on their suitability for VANET, we implemented them by using SUMO and NS3 as simulation tools applied on a real street map of Oujda city. We have extracted the used map from Open Street Map (OSM). Finally, we present our future works used for optimizing the greedy forwarding technique that is adopted by some position-based routing protocols in VANETs. Our suggested technique is based on the angle direction and three other important parameters of the relaying node.

1. Introduction

In this paper, we review and provide an extension version of previous work originally presented in the 5th International IEEE Congress on Information Sciences and Technology (IEEECiST'18) [1], in addition to some new enhancements. It is about routing protocols in Vehicular Ad-hoc Networks.

Vehicular networks have become the heart of the Intelligent Transport Systems or ITS in short, to improve the road traffic safety and to make a comfortable driving environment for drivers and passengers. In order to establish a communication in VANET, there is no need to any pre-existing infrastructure; this kind of networks consist of a set of vehicles containing wireless transceivers, named On-Board Unit (OBU). These OBUs help vehicles to exchange information between each other.

Compared to other sub-category of Mobile Ad-Hoc Networks, vehicular networks have many exclusive and specific features, such as diverse communication environments. Actually, in the rural environment, the issue of sparse connectivity is high and the traffic condition is simple; while for the city environment, the density and the number of obstacles (such as building trees and others) are high, which make the traffic condition very complex. Besides, the high speed of vehicles could be the main cause of the high dynamic topology and the regularly link breakage issue as the second and third features of VANETS. The last one is the variable network density, which is caused by the fact that the traffic density varies. This later depends on the type of environments and the moment of the day, which is very high in rush hour and medium or low in the other time.

Due to the above-mentioned features of vehicular networks, the proposition of an effective routing technique remains even more challenging to overcome. In this paper, we give our proposed taxonomy of routing techniques in vehicular networks [2] into routing based on network's topology, vehicle's position, multicast technique and broadcast technique as is shown in Figure 1.

The main objective behind the presentation of this taxonomy is to specify the advantages and disadvantages of each routing protocols' class, then declare the most convenient categories to VANETs. Furthermore, this declaration will be proved by a comparison study that is provided under two simulation tools. We also introduce our future work in which we suggest a new strategy

*Corresponding Author: Amina BENGAG, bengag.amina@gmail.com

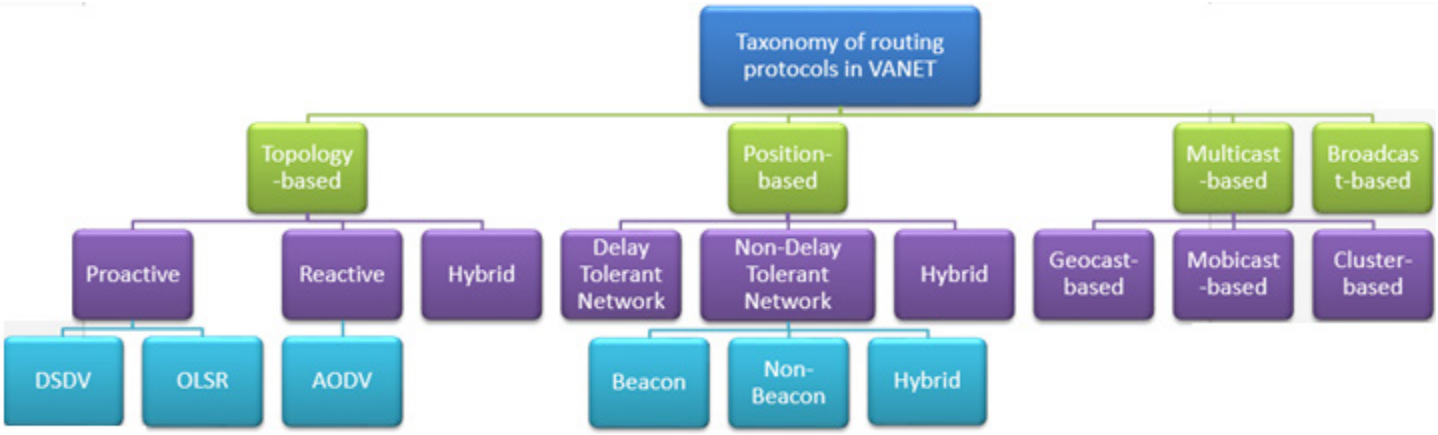


Figure 1 Taxonomy of VANETS routing protocols

to enhance the high-performed routing protocols, which belong to the category based on vehicles' position. This new strategy proved through theoretical analysis, is based on four parameters like velocity, the direction of vehicles, the density and the vehicles' position.

The rest of this paper is organized as follows: in section II, we outline the related works. In section III, we provide a detailed categorization of VANET's routing protocols and discuss their benefits and issues [2]. Section IV discusses in detail a comparison study of some routing techniques, based on the network topology and others on the physical position of vehicles by utilizing two simulation tools like NS3 as a network simulator and SUMO [1] as a road traffic simulator. Then, in section V, we introduce our current work and the proposed technique to optimize the Greedy forwarding strategy in GPSR protocol. Finally, in section VI, we conclude the paper and we present some our future works.

2. Previous work

To select the best-suited protocol to Vehicular Networks, many studies have been conducted to evaluate and examine the performance of certain routing mechanism based on simulation tools. In Table 1, we present some of the previous works related to the mentioned above study. It could be concluded from the table that all works is based on ns-2 or ns-3 as networks simulators, while some of them use also a specific road traffic simulators such as SUMO. However, in this work we provide a new study distinguish from those cited in Table I in many sides.

Basically, our work in this paper, focused on modeling the real details of VANETs network movement to examine the performance of certain routing protocols: DSDV, AODV, GPSR, OLSR, and GPCR. We extracted a part of a real road map from open street map (OSM) of a large street in Oujda city (Hay al-Quds), instead of considering a grid map or a random map. This map was made to be as an input file in SUMO in order to create mobility and traffic files; those files will be used as an entry in NS3 for further analysis of network performance.

3. Taxonomy of Routing Protocols in VANETs

The regular link breakage issue of vehicular networks caused by the speed of vehicles that lead to the high dynamic topology in these networks; the mentioned problems make the task of transmitting data packet in vehicular networks a very challenging one. Mainly, routing protocols in VANET can be classified into four categories [2] for V2V communication: based on the network's topology, based on the vehicles' position, based on multicast technique and finally routing based on broadcast technique.

Table 1 Related Works

Title and Reference	Simulators	Routing Protocols	Results
"Performance evaluation of olsr and aodv protocols in a vanet crossroad scenario"[3]	NS-3 CAVENET	OLSR, AODV	OLSR is the best one
"Comparative analysis of various routing protocols in vanet"[4]	NS-2.34	AOMDV, AODV, DSDV, DSR	The best performance is gotten in case of DSR for throughput and E2ED. Whereas, AOMDV and AODV perform better in terms of PDR
"Performance of routing protocols for vanets: A realistic analysis format" [5]	NS-3 and VANETMobiSim	AODV, OLSR	OLSR has the best performance

"Scenario based performance analysis of aodv and gpsr routing protocols in a vanet"[6]	NS-2.35 and VANETMobi Sim	AODV, GPSR	AODV outperforms GPSR in case of PDR metric and GPSR behaves better for E2ED
"Performance comparison of position based routing protocols for vanets"[7]	NS-2 and VANETMobi Sim	GPSR, GPCR	GPCR is the best one
"Performance Evaluation of OLSR and AODV in VANET Cloud Computing Using Fading Model with SUMO and NS3"[8]	NS-3 and SUMO	AODV, OLSR	Almost the same performance for both protocols

3.1. Routing based on the network topology

In this category, picking an itinerary between the source and the target node is based the existing link information in the network. The collected information could be done before any need of routing by the vehicle by utilizing proactive technique that is known by the table-driven technique. Moreover, for reactive technique, the information is searched when we need to transmit data; this procedure is also known by on-demand technique, in which the step of finding or maintaining an itinerary from source vehicle to destination is essential before any transmission.

In proactive technique, nodes collect and maintain updated routing data beforehand by periodically distributing routing tables throughout the network to all the nodes in its transmission range. The control messages in proactive routing are broadcasted continuously even if there is no data flow to send. Hence, the routing overhead is very high and the PDR is low consequently. The benefit of this category of routing protocols is that nodes can found routing information easily. DSDV, OLSR are two examples of proactive routing protocols.

In reactive technique, nodes start searching for a route only if they need to transmit data. Hence, they flood the network with Route Request packets (RREQ). AODV and DSR are two examples from this category. There is also one more type called hybrid protocol that associates reactive and proactive techniques, which make them more efficient. Zone Routing protocol (ZRP) is an example of such routing technique [9].

3.2. Routing based on vehicle's Position [10]

This class of routing protocols is also known as geographic protocols. They require the availability of a positioning system that provide the physical location of the participating vehicles in the relay selection process. The utility of this physical position information is identifying and making decision to select the next hop to route packet to the last destination. This technique reduces the routing overhead issue because it does not require any routing tables or exchange any information between them to establish routes. Here, it is not essential to identify the whole route to deliver the data packets that reduce the end-to-end delay. Based on several previous works, we can divide the routing based on vehicle's position into three sub-categories: routing adapted to Delay Tolerant Network (DTN), routing adapted to Non-Delay Tolerant Network (Non-DTN) and hybrid position-based routing.

Routing protocols for DTN: this category of routing techniques is considered as an efficient one for networks that have some specific features, such as high link breakage, limited bandwidth, power constraint and large scale. Those problems could be reduced by using a carry-and-forward technique, in which the transmitted vehicle stores the packet until it finds another vehicle moves into its vicinity. Vehicle-Assisted Data Delivery (VADD) [11] is an example of such routing technique.

Routing for Non-DTN: those protocols are characterized by the regularly link breakage issue that cause the lack of connectivity and a lack of instant end-to-end routes. The greedy approach is the process adopted by this routing to look for the next relaying vehicle. Actually, by utilizing this strategy, the packet is transmitted to the nearest neighbor of the destination. While the transmission failure could happen in anytime especially when a node finds no neighbors nearer to the destination than itself. In non-DTN protocols, we can find three sub-class; routing based on beacon that use a periodic hello message, others are beaconless, and hybrid protocols. GPSR and GPCR are an example of such routing technique considered as beacon-based protocols.

Hybrid position-based Routing protocols: this category of routing techniques could be considered as a combination of the two previous mentioned techniques. The main advantages of this category is reducing the control overhead and resolving the frequently link breakage problem in VANET. GeoDTN+Nav [12] is an example of such routing technique.

3.3. Routing based on multicast technique [13]

In routing based on multicast technique, the data packet is forwarded from a single source to a specific group of vehicles by using multi-hop communication technique. This protocol could be divided into three categories: protocols based on the geocast technique, based on Mobicast technique and those based on clusters.

- Routing based on Clusters [14] here, the vehicles share the same features such as the direction of mobility with almost the same velocity can form a cluster, and in each one a node will be elected to manage the communication into each cluster and between those clusters. This node is called a cluster head (CH). Cluster based Routing (CBR) [15] is an example of this category.

- Geocast-based routing protocols [16] here, data is forwarded from a single vehicle as source to vehicles belonging to multicast group or zone of relevance (ZOR) based on the physical position of vehicles. One of the most known routing in this category is Inter-Vehicule Geocast (IVG) [17].
- Routing based on Mobicast technique [18] it is a kind of geocast-based protocol that are based in time as another factor in addition to the space. Actually, by using this method, data packets will be transmitted to all vehicles situated in certain geographic area at a specific time t so called ZORt. One of the most popular routing in this class is Mobicast routing protocol with carry and forward [18].

3.4. Routing based on broadcast technique [19]

In the broadcast technique, the data packets will be transmitted to reach the whole vehicles in the network. This technique is usually used in vehicular networks to share some important information between vehicles, such as road conditions, traffic, weather and emergency event. Based on several previous works [20], [21], we can divided this class of routing protocols into three categories as shown in Figure 2. These categories could have several common properties such as information type, information source, broadcast decision and orthogonal enhancements.

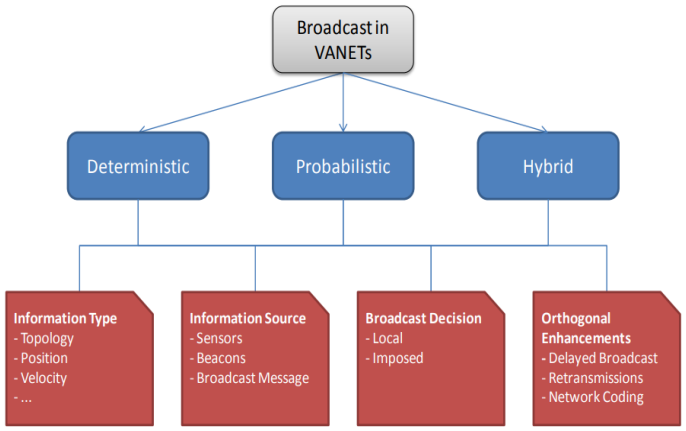
The three classes of routing protocols based on broadcast techniques are:

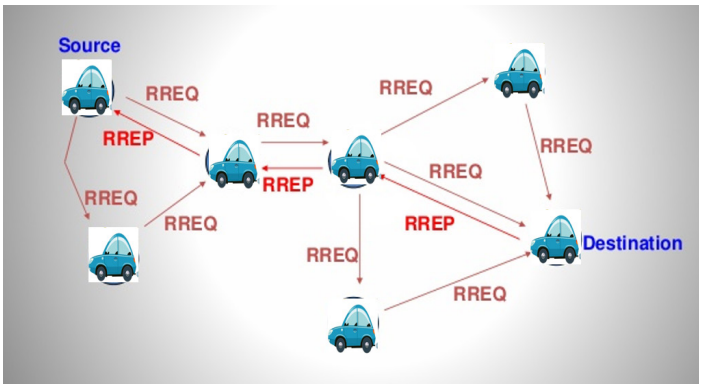
- The deterministic that is based on the topology of the network, the position and the velocity of the vehicles to rebroadcast the data packets.
- The probabilistic in which a vehicle will rebroadcast a data packet with a certain probability fixed in priory or adapted periodically.
- The hybrid broadcast that is a combination of deterministic and probabilistic techniques. In fact, the hybrid broadcast technique uses the deterministic technique to compute the most convenient broadcast probabilities.

Distribution Vehicular broadCAST (DV-CAST) [22], Distribution-Adaptive Distance with Channel Quality (DADCQ) protocol [23], Link-based Distributed Multi-hop Broadcast (LDMB) [24] and Optimized Position-based Gossiping (OPbG) protocol [25] are some examples of this category.

4. Comparison of routing based on network's topology and vehicle's position

This section reviews in detail the evaluated routing protocols that includes routing based on network's topology and others based on vehicle's position.





4.1. Routing based on network's topology

- AODV routing protocol. AODV [26] could be classified as unicast or multicast routing, the process adopted by this routing is that each paths are produced only on request. In fact, when a vehicle desires to transmit a packet by using the route discovery mechanism as Figure 3 shows, only the used routes are kept. In this case, the route discovery process causes significant delays before every data transmission to look for the path. AODV is also known by its capacity to diminish broadcasts, transmission latency and routing overhead. However, AODV suffers from the high end-to-end delay resulted by the route discovery process before every data transmission. Hence, the high E2ED is not suitable to vehicular networks in case of a crucial or dangerous information.
- OLSR routing protocol. In OLSR [27], each node constructs a global view of the network's topology by distributing after each specific time its routing table. In this routing technique, the concept of Multi point Relays (MPR) is utilized to relay control traffic. The MPR technique decreases the routing overhead issue by selecting only specific nodes to retransmit the control messages, which aids in optimizing the broadcast technique.

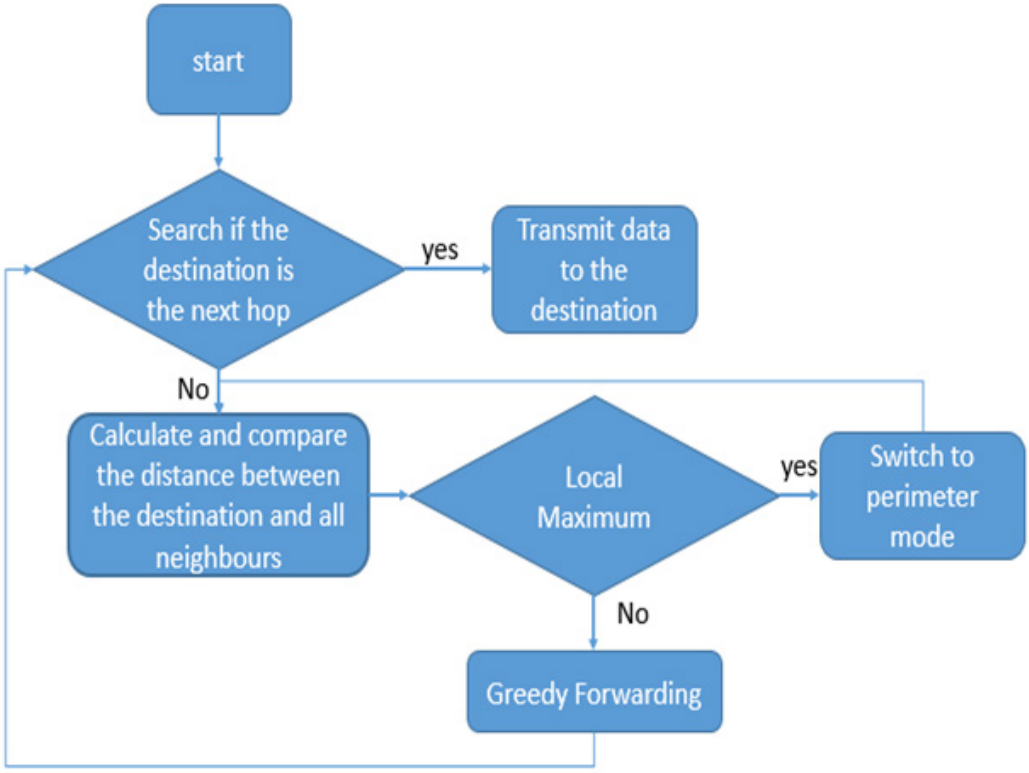


Figure 4 Flow diagram of route maintenance in GPSR

- DSDV routing protocol. In DSDV [28], [29], the update of routing tables of each nodes is done periodically to make available information about paths to each destination in the network at any time, even if the paths are at this time unused. In spite of the benefits of DSDV such as simplicity, loop free and no delay time caused by the discovery of the road technique. The DSDV still suffer from the high routing overhead caused by the periodic updates.

4.2. Routing based on vehicle's position

- GPSR [30] routing protocol. This routing protocol requires the availability of the physical location information of participating nodes using a location service like a GPS. In GPSR, the greedy forwarding (GF) technique is used by each relaying vehicle in order transmit data. In case of the local maximum problem, this technique fails; hence, the perimeter forwarding technique is applied to select another node through which a packet will travel. GPSR has many benefits: to transmit a data packet, a node necessities to memorize only the physical location of one hop neighbor. However, because of the features of high speed of vehicles, stale information of the neighbors' physical location are frequently contained in the sending nodes neighbor table. To be more clear, the strategy of GPSR is explained in Figure 4.
- GPCR routing protocol. Greedy Perimeter Coordinator Routing [31] protocol is an enhanced version of GPSR that uses a restricted greedy forwarding technique and a junction-based repair strategy in which routing decisions are made. Therefore, all forwarded packets must be sent to a vehicle located at an intersection or a junction; then this vehicle will transmit those packets to the closest neighbor of the recipient. GPCR has a great benefit that reduce the routing overhead problem and the end-to-end delay. Indeed, in the transmission process the election of the next hop is depends only on junction vehicles, which does not require any global or external information. Nevertheless, GPCR has some drawbacks that badly affect its performance. Actually, this technique assumes that there is at all times at least one vehicle at an intersection, which is not always the case.

4.3. Simulation and comparison

To simulate and evaluate the performance of the above-mentioned routing protocols in VANETs, it would be very expensive to set up a network to test certain criteria. To remedy this problem, we will use network and traffic simulators. In this project, we realize a scenario where we test the performance of routing protocols by successively increasing the number of vehicles in the network, while the other network parameters are fixed.

Basically, our work was evaluated by computing the rate of three metrics as the received packet number, the end-to-end delay and the routing overhead. To ensure this task, a simulation was carried out under the network simulator NS3 [32] version 25 in combination with the road traffic simulator SUMO 0.25 [33] by using the operating system Ubuntu 14.04. We have also chosen the Open Street Map (OSM) [34] to download a real-world urban map for a part of Oujda city (Hay al Quds) in Morocco; after that,

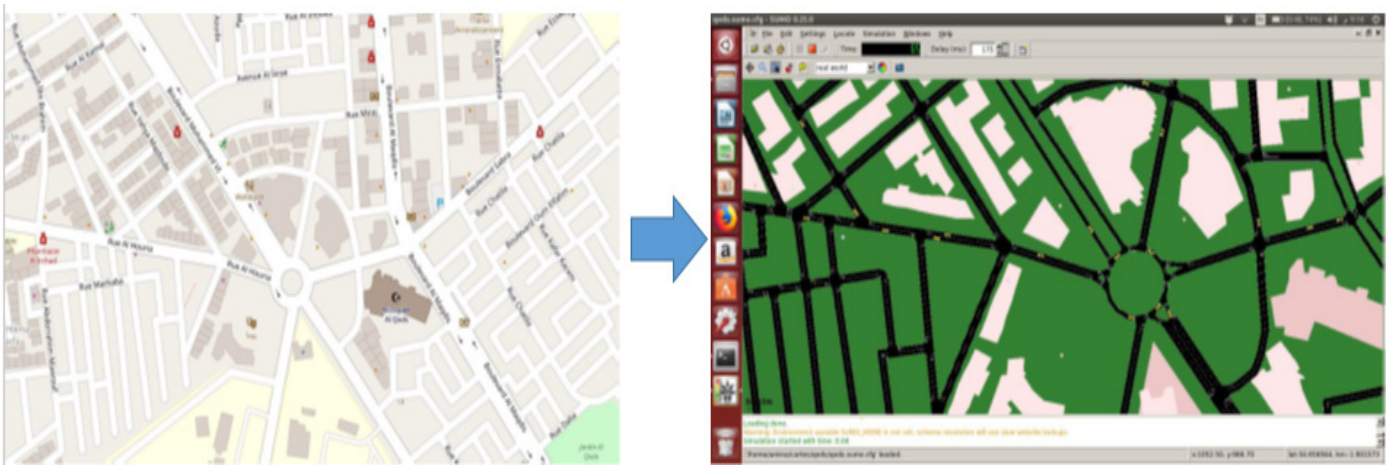


Figure 5 Hay Alquds Oujda map from OSM to SUMO

we have used SUMO to generate the traffic scenario as shown in Figure *5*.

In the used scenario, the number of nodes connected in the network varies. The simulation configuration is shown in Table *1*.

Table 1 Setting of experimental simulation parameters

Parameters	**Measures**
Routing Protocols	AODV, OLSR, DSDV, GPSR, GPCR
Simulation Object	Hay-ElQuds Oujda city
Number of nodes	20,30, 40, 50, 60, 70, 80, 90
Vehicles speed	Max : 20 m/s
Simulation time	100 s
Simulation area	1.7 * 1.5 km^2
Packet size	512 bytes
Data type	Constant Bit Rate (CBR)
Transport protocol	UDP
Propagation model	Two Ray Ground
Mac protocol	IEEE 802.11p
Channel	Channel/WirelessChannel
Network interface	Phy/WirelessPhy
Transmission range	145m
Transmission power	20 dBm

- Packet Delivery Ratio (PDR)
  According to the Figure 6, the results obtained confirm that DSDV and OLSR protocols have the best values of PDR. These values are increased when the vehicle's number increases, while they are reduced in case of AODV. Moreover, the curve stays stable For GPCR and GPSR protocols but with a moderate values of PDR. In fact, the huge number of radio obstacles in the proposed scenario are the main cause of this weak. Those obstacles increase the possibility of getting into the local maximum's trouble as explained in section IV. In addition, the PDR for GPSR and GPCR may be negatively affected because of the perimeter mode technique where the number of hops become high. For OLSR protocol, the good values obtained may be thanks to the use of Multi-Point Relay technique, which has a positive effect in case of dense and large networks.

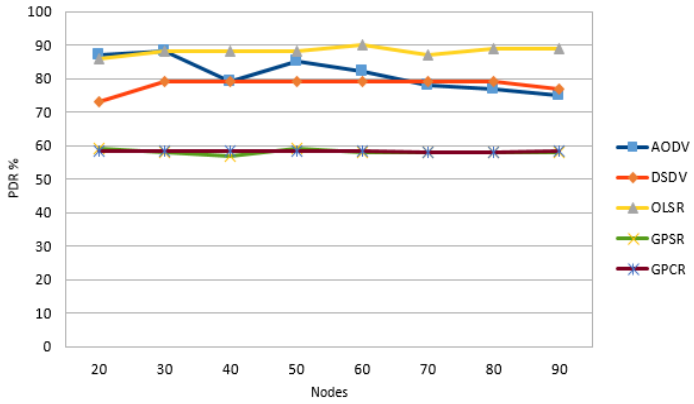


Figure 6 Packet Delivery Ratio Vs number of vehicles

- Average End-to-end Delay
  Average End-to-end Delay or E2ED in short denotes the required time for transmitting a packet from the source node to the destination per the number of successfully transported and received packets. The protocol that has an average E2ED as low as possible is the good and the preferred one. To calculate the E2ED we used the equation (2).

$$E2ED = \frac{\sum_1^n receptionTime - SendTime}{ReceivedPackets} \quad (2)$$

From the graph showed in Figure 7, we approve that the OLSR protocol has the maximum values followed by DSDV. Conversely, GPSR and GPCR have the lowest values, which are almost the same for all the number of nodes. The achieved results for GPSR and GPCR could be thanks to the use of the greedy forwarding technique. By using this technique, the number of hops are reduced so as the necessary time for transmitting data packets until the endpoint vehicle. It is also clear, that AODV has the best values in case of the E2ED compared to DSDV and OLSR for low node density. However, for the high node density, the E2ED for AODV are high because of the use of the route discovery technique that introduces additional delay.

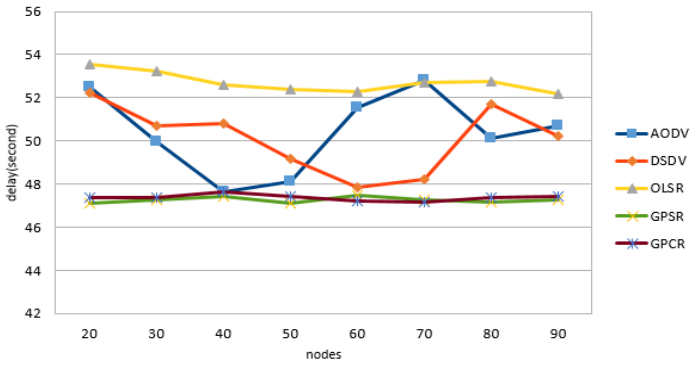


Figure 7 E2ED vs number of vehicles

- Throughput:
  The graph in Figure 8 proves that OLSR has the best and highest values of throughput up to 18kbps. In fact, the use of Multi point relay technique reduces the number of relaying nodes so as the routing overhead that helps in increasing the throughput. In case of DSDV and AODV, the throughput decreases as the density increases; this could be explained by the fact that DSDV broadcasts the entire routing table after fixed time interval, which introduces an additional overhead and affects its throughput. While GPSR and GPCR have the lowest throughput compared to the other three protocols.
- Routing Overhead:
  The results obtained in Figure 9 show that the best results are in case of GPSR and GPCR protocols that have the smallest values of routing overhead. This result could be explained by the fact that GPSR and GPCR are based on the physical location of the destination that is contained in the routed packet so that the retransmitted node should not need to use the location service for a second time, which reduces the overhead. The highest control overhead is for AODV, as it broadcasts a large number of control message packets to maintain the route.

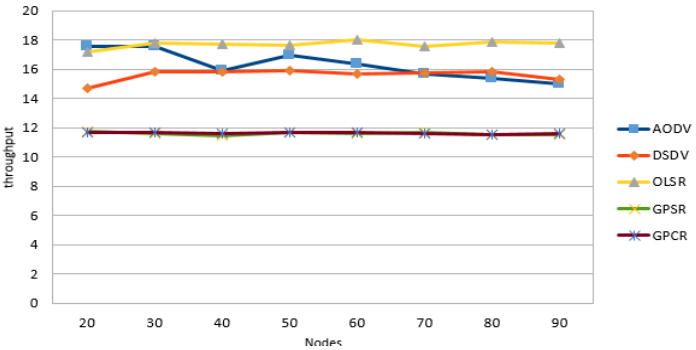


Figure 8 the throughput vs number of vehicles

5. Work in progress

In our proposed scenario, the performance evaluation above proves that OLSR has the best results in case of PDR and throughput. At the same time, GPSR and GPCR have the best results in case of overhead and E2ED. Indeed, the last two protocols are based on the destination's position that is carried in the packet, so the transferring vehicles do not need to use the location service for a second time and thus reduce the issue of routing overhead.

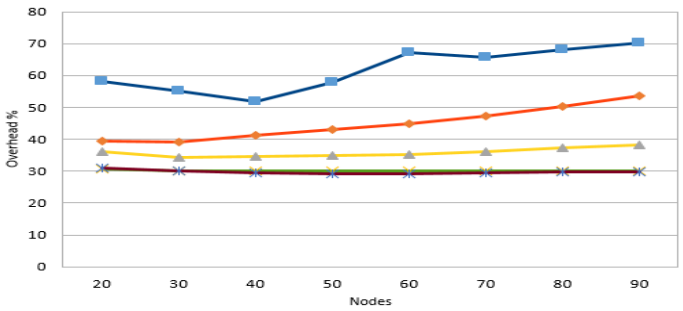


Figure 9 routing overhead vs number of vehicles

GPSR and GPCR have a moderate amount of PDR and Throughput. Hence, to optimize the ratio of the mentioned parameters, we proposed a new algorithm that improves the classical technique of greedy forwarding. This novel technique is used to select the best relaying vehicle candidate by taking into consideration four main metrics: the angle direction, the speed variation, the density of each nodes in addition to the distance between the next-hop candidate vehicle and the destination vehicle.

Our main contribution consists in suggesting a new greedy forwarding strategy based on a simple weighted function (WF), which uses the four last mentioned metrics. Then, an improved GPSR and GPCR protocols will be provided based on our proposed strategy. More specifically, the chosen of the best relaying node candidate is done by selecting the neighbor node whose header angle is closest to the destination node. Therefore, the connection between the source and the destination consists of a series of nodes that move approximately in the same direction. The main goal of this concept is to maintain the connection between the nodes as long as possible, in order to reduce the problem of link breakage. This is in contrast to the simple greedy forwarding, which only considers the distance between the next hop and the destination and does not take into account the connection lifetime. The new technique also could enhance the stability of the routing link, which mitigates the influence of high dynamic topology in VANETs networks.

In GPSR and GPCR, without considering moving direction of vehicles while looking for the next optimal hop, the routing techniques could produce an erroneous decision to transmit a data packet, which could result in a high packet loss issue. In addition, the regularly link breakage problem may also disturbs the quantity of delivered packets, particularly in highway environment where the speed is very high. Basically, when a node receives a Hello packet to up-to-date its neighbor's list, it computes the best next hop and then forwards packet. Thus, during this process, the selected neighbor could be outside the transmission range and the packet cannot reach it due to the high network's mobility. In order to solve those issues that lead to performances degradation, we

thought to use other parameters as mentioned above like the speed, the density of neighbors and the distance between all participating nodes and the destination during routing operation in addition to the moving direction. Furthermore, a simple weighted function (WF) will be applied on each relaying candidate vehicle, then the relaying candidate node that has the lowest value of the WF will be selected.

The proposed technique resolves mainly three problems in vehicular networks, which could be proved through theoretical analysis:

1. The problem of void area is resolved by taking into account the density (number of neighbors) of the selecting node. In Figure 10, we present the density parameter or the number of neighbors of a selected node. Basically, vehicle A has the vehicles B to F as neighbors because they belong to its transmission range.

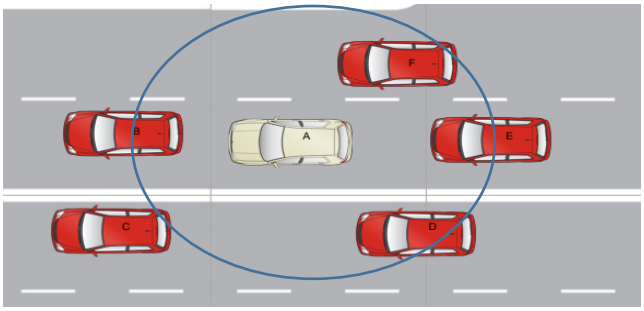


Figure 10 neighbors of the node A (density)

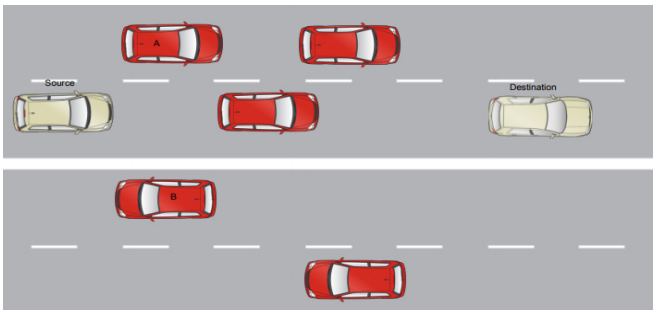


Figure 11 void area problem

Indeed, by using the proposed technique the node that has the big number of neighbors increases the probability to be chosen as the next hop. Figure 11 presents an example in which the source will choose A as the next hop because it has three neighbors while B has no neighbors.

2. The problem of direction of nodes in VANETs will be fixed by taking into account the angle direction calculated between the next hop and the destination node. As the Figure 12 shows, the source node will choose A as the next hop by applying our proposed technique because it has almost the same direction as the destination. The angle direction between the next hop and the destination is calculated by using our proposed formula (3).

$$\varphi_{id} = \cos^{-1}\frac{((iVelocity.x*dVelocity.x)+(iVelocity.y*dVelocity.y))}{(\sqrt{(iVelocity.x^2+dVelocity.x^2)}*\sqrt{(iVelocity.y^2+dVelocity.y^2)})} \quad (3)$$

Where iVelocity is the velocity of the next hop candidate and dVelocity is the destination velocity. The rational between the concepts of the angle direction is to maintain the connection between vehicles as long as possible by choosing the small value of all calculated φ_id.

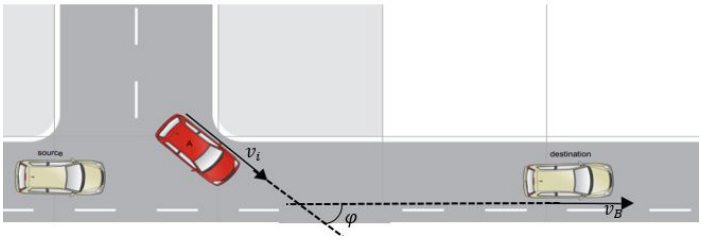


Figure 12 the angle direction φ

6. Conclusion

In this paper, we have provided a taxonomy of routing protocols in VANETS. Then we have given a quantitative analysis of the most known routing protocols for vehicular networks, using two simulation tools SUMO as a road traffic simulator and NS-3 as a network simulator, to determine the benefits and challenges of each algorithm.

From the results of the simulation, we have picked up the feebleness of GPSR and GPCR in terms of PDR and throughput. Based on the given results and to enhance the performance of those protocols, we have proposed a technique that combines the simple greedy forwarding technique and some other parameters selected carefully such as the direction, the speed and the density to choose the next relaying vehicle.

As a future work, we plan to develop our proposed technique to enhance GPSR and GPCR then prove their performance compared to the original GPSR and GPCR.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research is performed inside the MATSI Lab., ESTO, University Mohammed First, Oujda (Morocco).

References

[1] B. Amina and E. Mohamed, "Performance Evaluation of VANETs Routing Protocols Using SUMO and NS3," in *Colloquium in Information Science and Technology, CIST*, 2018, vol. 2018-Octob, pp. 525–530.

[2] A. Bengag and M. El Boukhari, "Classification and comparison of routing protocols in VANETs," in *2018 International Conference on Intelligent Systems and Computer Vision, ISCV 2018*, 2018, vol. 2018-May.

[3] E. Spaho, M. Ikeda, L. Barolli, and F. Xhafa, "Performance comparison of OLSR and AODV protocols in a VANET crossroad scenario," *Lect. Notes Electr. Eng.*, vol. 253 LNEE, pp. 37–45, 2013.

[4] S. Singh, P. Kumari, and S. Agrawal, "Comparative Analysis of Various Routing Protocols in VANET," *2015 Fifth Int. Conf. Adv. Comput. Commun. Technol.*, pp. 315–319, 2015.

[5] B. V. Coutinho, E. C. G. Wille, and H. I. Del Monego, "Performance of routing protocols for VANETs," *Proc. 9th Int. Conf. Ubiquitous Inf. Manag. Commun. - IMCOM '15*, pp. 1–6, 2015.

[6] R. Bala and C. R. Krishna, "Scenario Based Performance Analysis of AODV and GPSR Routing Protocols in a VANET," in *2015 IEEE International Conference on Computational Intelligence & Communication Technology*, 2015, pp. 432–437.

[7] A. Ashtaiwi and S. Abdusadik, "Performance Comparison of Position based Routing Protocols for VANETs," pp. 78–84, 2017.

[8] G. Sallam and A. Mahmoud, "Performance Evaluation of OLSR and AODV in VANET Cloud Computing Using Fading Model with SUMO and NS3," *2015 Int. Conf. Cloud Comput. ICCC 2015*, pp. 1–5, 2015.

[9] Z. J. Haas, M. R. Pearlman, and P. Samar, "The Zone Routing Protocol (ZRP) for Ad Hoc Networks," *draftietfmanetzonezrp02 txt*. p. 11, 2002.

[10] J. Liu, J. Wan, Q. Wang, P. Deng, K. Zhou, and Y. Qiao, "A survey on position-based routing for vehicular ad hoc networks," *Telecommun. Syst.*, vol. 62, no. 1, pp. 15–30, 2016.

[11] Jing Zhao and Guohong Cao, "VADD: Vehicle-Assisted Data Delivery in Vehicular <emphasis emphasistype="italic">Ad Hoc</emphasis> Networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 3, pp. 1910–1922, May 2008.

[12] P.-C. Cheng, K. C. Lee, M. Gerla, and J. Härri, "GeoDTN+Nav: Geographic DTN Routing with Navigator Prediction for Urban Vehicular Environments," *Mob. Networks Appl.*, vol. 15, no. 1, pp. 61–82, Feb. 2010.

[13] Y. Lin, Y. Chen, and S. Lee, "Routing Protocols in Vehicular Ad Hoc Networks A Survey and Future Perspectives.," *J. Inf. Sci. ...*, vol. 932, no. c, pp. 913–932, 2010.

[14] B. Marzak, H. Toumi, E. Benlahmar, and M. Talea, "Analysing Clustering Routing Protocols Performance for Vehicular Networks," *Proc. 2nd Int. Conf. Big Data, Cloud Appl. - BDCA'17*, pp. 1–7, 2017.

[15] Y. L. Y. Luo, W. Z. W. Zhang, and Y. H. Y. Hu, "A New Cluster Based Routing Protocol for VANET," *Networks Secur. Wirel. Commun. Trust. Comput. NSWCTC 2010 Second Int. Conf.*, vol. 1, pp. 176–180, 2010.

[16] S. Allal, S. Allal, and S. Boudjit, "Geocast Routing Protocols for VANETs : Survey and Geometry-Driven Scheme Proposal Geocast Routing Protocols for VANETs : Survey and Geometry-Driven Scheme Proposal," no. February 2013, 2015.

[17] A. Bachir and A. Benslimane, "A multicast protocol in ad hoc networks inter-vehicle geocast," *IEEE Semiannu. Veh. Technol. Conf.*, vol. 4, pp. 2456–2460, 2003.

[18] Y. S. Chen and Y. W. Lin, "A mobicast routing protocol with carry-and-forward in vehicular ad hoc networks," *Int. J. Commun. Syst.*, vol. 27, no. 10, pp. 1416–1440, 2014.

[19] S. Zeadally, R. Hunt, Y.-S. Chen, A. Irwin, and A. Hassan, "Vehicular ad hoc networks (VANETS): status, results, and challenges," *Telecommun. Syst.*, vol. 50, no. 4, pp. 217–241, 2012.

[20] B. Williams and T. Camp, "Comparision of broadcasting techniques," pp. 194–205, 2002.

[21] B. Bako, "Efficient Information Dissemination in VANETs," 2016.

[22] O. Tonguz, N. Wisitpongphan, F. Bai, P. Mudalige, and V. Sadekar, "Broadcasting in VANET," in *2007 Mobile Networking for Vehicular Environments, MOVE*, 2007, pp. 7–12.

[23] M. Slavik and I. Mahgoub, "Spatial distribution and channel quality adaptive protocol for multihop wireless broadcast routing in VANET," *IEEE Trans. Mob. Comput.*, 2013.

[24] Q. Yang and L. Shen, "A Multi-Hop Broadcast scheme for propagation of emergency messages in VANET," in *International Conference on Communication Technology Proceedings, ICCT*, 2010.

[25] B. Bako, E. Schoch, F. Kargl, and M. Weber, "Optimized position based gossiping in VANETs," in *IEEE Vehicular Technology Conference*, 2008.

[26] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *Proceedings - WMCSA'99: 2nd IEEE Workshop on Mobile Computing Systems and Applications*, 1999, pp. 90–100.

[27] T. Clausen and P. Jacquet, "Optimized Link State Routing Protocol (OLSR)," *Ietf Rfc3626*, p. 75, 2003.

[28] C. E. Perkins and P. Bhagwat, "Highly dynamic Destination-Sequenced Distance-Vector routing (DSDV) for mobile computers," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 24, no. 4, pp. 234–244, 1994.

[29] G. He, "Destination-sequenced distance vector (DSDV) protocol," *Netw. Lab. Helsinki Univ. Technol.*, pp. 1–9, 2002.

[30] B. Karp and H. Kung, "GPSR: Greedy Perimeter Stateless Routing for wireless networks," *ACM MobiCom*, no. MobiCom, pp. 243–254, 2000.

[31] C. Lochert, M. Mauve, H. Füßler, and H. Hartenstein, "Geographic routing in city scenarios," *ACM SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 9, no. 1, p. 69, 2005.

[32] R. Manual *et al.*, "ns-3 Tutorial," *System*, no. January, pp. 1–46, 2011.

[33] M. Behrisch, L. Bieker, J. Erdmann, and D. Krajzewicz, "SUMO – Simulation of Urban MObility," *Iaria*, no. c, pp. 55–60, 2011.

[34] OpenStreetMap Foundation, "OpenStreetMap," *Open Database License (ODbL)*. 2013.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 80-86 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Attacks Classification and a Novel IDS for Detecting Jamming Attack in WBAN

Asmae Bengag*, Amina Bengag, Omar Moussaoui

Applied Mathematics, Signal Processing and Computer Science Laboratory, ESTO, University Mohamed 1er Oujda, Morocco

ARTICLE INFO

Article history:
Received: 01 January, 2020
Accepted: 16 February, 2020
Online: 09 March, 2020

Keywords:
WBAN
Medical Sensors
Jamming attack
IDS
Parameters Network
ZigBee
OMNET++
Castalia 3

ABSTRACT

Wireless Body Area Network (WBAN) aims to monitor patient's health remotely, by using mini medical sensors that are attached on the human body to collect important data via the wireless network. However, this type of communication is very vulnerable to various types of attacks, poses serious problems to the individual's life who wears the nodes. In this paper, we present a new classification of the most dangerous attacks based on different criteria, which gives us a clear vision of how attacks affect a WBAN system. Moreover, this classification will help us to specify the strength and the weakness of each attack in order to facilitate the development of a new intrusion detection system (IDS). In the second part of this work, we develop a novel IDS for detecting three types of jamming attacks in WBAN. The proposed methodology is based on the network parameters as an indicator to differentiate the normal case from the abnormal case like false alert or attack state. Through simulation analysis that was applied on Castalia platform by using OMNET++ as a simulator, proves that the proposed IDS have a great effect for detecting the presence of jamming attack in the network.

1. Introduction

In the last two decades, wireless body area network (WBAN) has attracted huge number of researchers. WBAN is a network consisted of several types of medical sensors; each one has its own debit and functionality. These sensors facilitate the supervision of the patient's health and intervene as quickly as possible in emergencies.

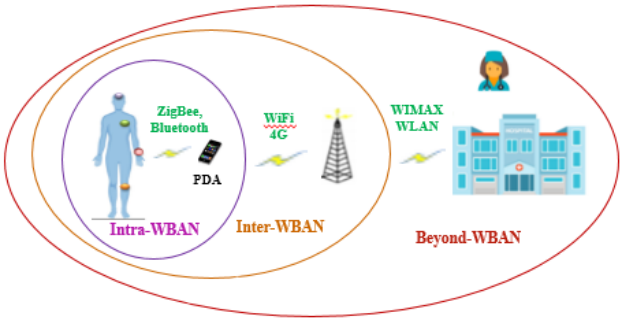


Figure 1 : Communication architecture of WBAN

The medical sensors are attached on the human body for collecting the biomedical parameters as the activity of muscles, brain or heart rate. Then, this information will be transmitted to the personal device assistant (PDA). Finally, the data will be received by the medical team via a wireless mode, as shown in Figure 1.

Actually, the communication is done by three levels [1]:

- Level 1: 'Intra-WBAN communication' that contains different devices especially medical sensors and the coordinator node or PDA. The communication between these nodes could be based on Bluetooth (802.15.1) or ZigBee (802.15.4). In fact, the main goal of PDA is to ensure a communication between sensors, and informs the final user as doctor or patient via an external gateway [2].

- Level 2: is named an 'Inter WBAN communication' that concerns a communication between the PDA and one or more access points (APs) via WiFi, to interconnect WBAN network with several networks [3].

- Level 3: called a 'Beyond-WBAN communication' in which the communication is between a various access points and medical server (hospital, nurse etc.) using for instance Internet or mobile networks.

However, the communication used in WBAN system is not reliable one hundred percent, because the wireless communication deployed on open radio frequencies with various attacks and different cases of anomalies. Furthermore, several attacks threat

*Asmae Bengag, asmaebengag@gmail.com

the security aspects like availability, which makes the medical nodes not able to send or transmit any information especially in the case of emergency. Besides, other attacks affect the integrity that can modify the data transmitted during the transmission and other ones threat the confidentiality aspect that could access sensitive data [4].

One of the most important challenges to secure the WBAN system is by using a robust mechanism to detect an attack in this kind of networks. Hence, we started our work by studying the impact of the most dangerous attacks in WBAN, by classifying them according to various criteria. Then, we focused on the Denial of Service (DoS) jamming attack that disturbs wireless communication by affecting the physical and MAC layers. In fact, we are developed a new IDS technique to detect jamming attacks that based on four important network parameters: packet delivery ratio (PDR), energy consumption amount (ECA), received signal strength indication (RSSI) and bad packet ratio (BPR).

The rest of this paper is organized as follows: in section II, we describe the previous different classification attacks works, and the various IDS mechanisms for detecting jamming on the network. Then in section III, we present our attacks classification in WBAN according to various measures. The steps of our IDS algorithm, the simulations scenarios and the results are discussed in section IV. Finally, our work is concluded in section V.

2. Related Work

In order to propose a new classification of attacks in WBAN, we are based on several previous works. In fact, each work is based on different measures and applied on specific technology such as MANET, WSN or WBAN.

In [5] authors proposed a classification attacks that applied in MANETs and based on six attributes: legitimacy of attacking node(external or internal node), the number of nodes participating as an attacker (singleton or collision attack), MANETs vulnerabilities utilized by the attack, the network resources exploited by the attacking node/s, the targeted victim and compromised security service of attack.

Another taxonomy proposed by Wu et al. in [6] that is applied in network and computer, which based on three dimensions: (i) the source of attack (local or remote); (ii) the techniques dimension that mean the all-possible techniques adopted by the attackers; and (iii) the results of attack. However, the weakness of this work is the less information about the type of target.

In [7], the authors present various attacks get into IEEE 802.15.4 and classifying them into three categories: (i) the radio jamming that is an attack created the problems of denial of service in the physical layer; (ii) the message manipulation attacks for injecting false information into the node network; and (iii) the last one is the steganography attacks. However, the weakness of this classification is the lack of other important measures as security aspects and position of attack in the network that will be presented in our classification.

In [8] authors classified attacks according MAC Layer basing on the functionality of the mac layer. They are presented two classification, the first one concerns the MAC layer as Guaranteed Time Slot (GTS) attack, and the second one based on MAC layer protocol rules.

The OSI model is important measure to identify problems of security in each layer; this work is proposed by Pooja et al. [9] and applied in MANET. Nonetheless, this classification is based on just one criterion and it is not enough to have a robust security solution.

All the above-mentioned classification has some limits that make them inappropriate to specify all types of attacks because they not consider some important metrics. Basically, in our classification, we are based on the important and clear metrics especially in WBAN as attack position basing on the communication architecture of WBAN and the attack functionality.

In this second part, we give an overview about the detection mechanisms that have been studied to detect jamming attack in a wireless network.

Fuzzy Inference System (FIS) is a technique proposed by Reyes et al. [4] that built in Matlab tool using the following metrics as input values: CCA, BPR, RSS and PDR. In order, detecting the link loss in wireless networks, these parameters will be used to calculate the values of the Jamming index (JI). The idea of this method is good, but it has a complicated calculation that makes sensor nodes consume more energy. In addition, it does not differentiate the types of jamming present in the network.

In [10] proposed a technique for identifying the following jamming types: constant, random, reactive and deceiving. The methodology is based on SS (Signal Strength) and PDR (Packet Delivery Ratio), to determine the normal cases from the jamming attack cases. However, the method could consider that the sensor is under a jamming attack but there is no attack in the network. For instance, in the case where the receptor sensor has a problem to respond with an ACK packet to the transmitter, we found that the transmitter node has a high SS level, and the PDR value is low. Hence, there is confusion between jamming scenario and affected node.

In [11], the author improved an IDS that includes Signal Strength (SS) and Packet Delivery Rate (PDR) to define the presence of jamming attack. On the other hand, they are used also Packet Send Rate (PSR), to distinguish different types of jamming attacks in WSN.

In [12] proposed jamming detection technique for WSN that called physical layer jamming identification using PDR and RSSI parameter. This method used only some nodes as monitor in the network that have residual energy. Nevertheless, for monitoring the entire network, the method needs to implement various monitor nodes.

By comparing our proposed methodology with the previous works, we are focused on four important network parameters PDR, ECA, RSSI and ECA that will be presented in the section 4. By using these parameters based on a simple technique that consume a less energy compared to the above-mentioned techniques. Moreover, these parameters have a good effect to differentiate the normal case and false alerts for the jamming attacks. Our intrusion detection system allows also us to identify three types of jamming

attacks (constant jamming, reactive jamming and deceptive jamming) that are applied in the WBAN system.

3. Classification of attacks in WBAN

In this section, we describe our classification that based on previous attack taxonomies and other metrics, which adopt in wireless body area networks. Our taxonomy is based on six important criteria as follow:

- Attack impacts: are the results or the effects of attack action in the network.
- Security aspects: for having a robust system, we must take into consideration the fundamental security requirements in the medical application [13], such as data integrity, availability, data confidentiality, data freshness and authentication.
- OSI model: we can also classify attacks by the layers of the OSI model.
- Communication architecture of the WBAN: attack could place in different levels of the architecture of WBAN, either in the first level that directly threatens the medical sensors, level 2 or level 3 as exposed in the Table 1.
- Attack position in the network: we can distinguish two types of attack position, as internal attack that is a part of the network action, or as an external attack that is not a part of the system [5], like sniffing and man in the middle.
- Interruption act: the attack could be active by modifying or damaging the transmitted message to the receiver, and disrupting the communication [9]. Further, the passive attack threats the confidentiality aspect by collecting the data in the network without perturbing the communication [5].

Thus, the Table 1 presents the different attacks classified according to our taxonomy that are cited above.

This attack classification facilitates us to understand and identify the operation or strategy of each attack in the network. Furthermore, it helps us to find a robust mechanism to detect jamming attack in WBAN that will present in the next section.

4. A novel detecting jamming attack

Jamming attack is among one of the attacks that blocks and disrupts the communication between nodes, by sending illegitimate signals, in order to make the system unavailable. In other words, this attack generates interference, which makes the medical sensors consume a lot of energy, and involves the collision between them [14].

As we present in the previous section, jamming attack threats the physical layer that transmit a signal to create interference. Besides, data link layer could be under jammer node that does not respect the mechanism of its MAC protocol.

Table 1 Attacks classification in WBAN system

Attacks	Attack impacts	Security aspects	OSI model	Communication architecture of the WBAN	Attack position in the network	Interruption act
Tampering	Extracting cryptographic keys from the captured node [15]	Confidentiality Integrity data	Physical	Intra-WBAN	External	Active
Jamming	Disruption of communication by sending radio waves at a same frequency in the wireless network.	Availability	Physical Data Link	Intra and Inter WBAN	External	Active
Sybil	Obstruct the operation of routing protocols by operating multiple fake identities.	Confidentiality Integrity data	Network Data link	Intra, Inter and Beyond WBAN	Internal / External	Active
Hello floods	Blocking communication by sending broadcasts HELLO packets with high transmission.	Availability Confidentiality	Network	Inter and Beyond WBAN	External	Active
Selective forwarding	Delete some packets and refuse the transmission of data.	Availability, Confidentiality	Network	Inter and Beyond WBAN	External	Active
Flooding	Exhaust the energy of sensor and saturate the network.	Availability	Network Transport	Inter and Beyond WBAN	External	Active

4.1. Proposed methodology

In this section, we present our proposed methodology for identifying and detecting the presence of jamming attacks in WBAN system.

Our proposed algorithm has the ability to detect the presence of a jamming attack, and then identify which type it is. In fact, there are many types of jamming attack, but in our work, we mainly focus on three of them, constant jamming, deceptive jamming and reactive jamming.

Indeed, there are a lot of cases that resemble as jamming attack in the network namely false alerts, for example problems caused by a low energy or collision in the nodes, which reduce the quality of the IDS. In order, to have good IDS defending a malicious activities jamming in WBAN system, we are based on four important network parameters as a performance metrics that are mentioned in Table *2*.

Table 2 : The network settings used

Parameters	Definition
PDR (Packet Delivery Ratio)	The authors of [10][16] defined the PDR as the ratio of the packets successfully sent by the node, to the total sent packets.
ECA (Energy Consumption Amount)	Amount of energy consumed in a specified time [4].
BPR (Bad Packet Ratio)	BPR is a ratio of the failed packets received by a node [16].
RSSI (Received Signal Strength Indication)	RSSI is defined as a power content of the received radio signal at the receiver [4].

The diagram in Figure 3 shows the essential algorithm steps of the proposed IDS, which describes how our mechanism defend the remaining or not of jamming, and differentiate which types are on WBAN. As first step, we assume that the WBAN system is operating normally, without jamming attack or any problems. Then, each receiver sensors medical are calculated the thresholds of the network parameters PDRth, ECAth, BPRth and RSSIth. After that, for observing if there are any problems on the WBAN, the initial parameters (PDRth, ECAth, BPRth and RSSIth) are compared with the current values.

For mentioning the presence of jammer node and specify its type, one of the conditions will be launched an alert, and we can conclude that the node is under jamming attack. For the constant jamming, could be generated interferences between nodes sensors, by sending continually a random bits or signals in the wireless network [4], and forces the legitimate nodes to stay in listening mode, which cause a higher energy consumption. That is why the PDR and BPR values are lower than the normal case, and the ECA and RSSI parameters are higher.

In the case where the legitimate node is under the deceptive jamming, the latter sends data regularly to the channel that resembles legal, in order to deceive each node. Hence, we could be found that the PDR value is lower than the PDRth, and for the BPR, RSSI and ECA values are higher than their thresholds.

Furthermore, to identify a reactive radio jammer when the legitimate node sends the packet RTS, then the attack starts to transmit data and makes the communication medium busy. Thence, the legitimate node will consume a modest rate of energy. However, the value of the PDR input parameter very lower, whereas the BPR and RSSI values are higher than the normal values and the packet delivery ratio is very lower.

4.2. Simulation and results

4.2.1. Simulation parameters

We have made two main scenarios, in the first one we are simulated a WBAN system in normal case without malicious node, the studied scenario covers three types of medical sensors (EMG, ECG and EEG) are attached on human body. More specifically, the wireless communication protocol used between legitimate sensors and PDA is ZigBee (802.15.4 MAC), and CC2420 Radio used as a radio model as shown in Figure 2.

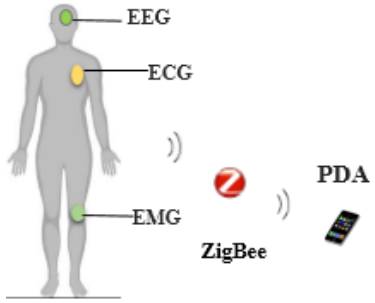


Figure 2: Simulation of WBAN in normal case

Simulation parameters, used in this case, are listed in Table 3.

Table 3 : Simulation parameters for WBAN without jamming

Used parameters	Values
Nodes	4 (3 medical sensors and 1 PDA)
Simulation time	300s (second)
MAC protocol type	ZigBee (Basic802154)
End simulation	End of the simulation

After that, we are simulated a WBAN system under a jammer node, as shown in Figure 4, in order to understand the impacts of jamming attack in the WBAN system and specially the MAC layer and the network parameters to compare with the normal case.

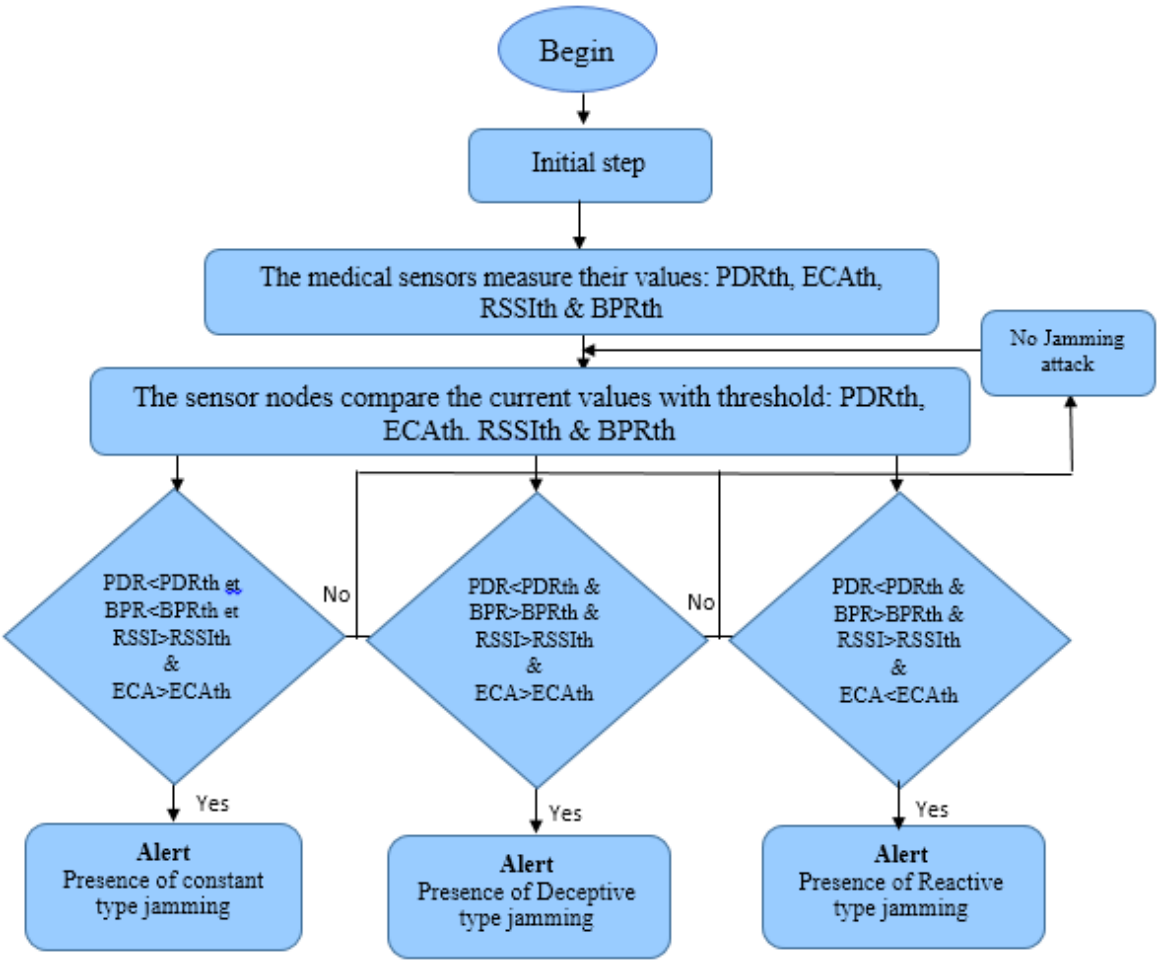


Figure 3 : Proposed mechanism for detecting different types of jamming attacks

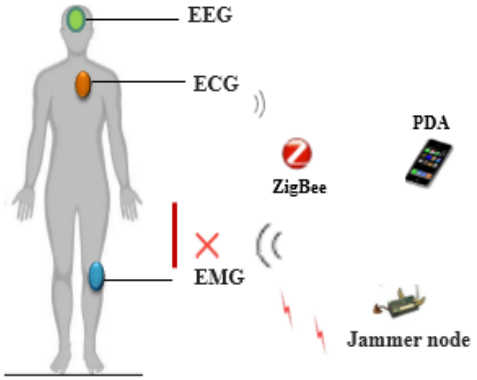


Figure 4: The scenario of WBAN with jamming attack

Table 4: Simulation parameters for jammer node

Used parameters	Values
Number of jammer nodes	1
Simulation time	300s (second)
MAC protocol	BypassMAC
Constant Data Payload	2000
End simulation	End of the simulation

As mentioned in Table 4, we are chosen the BypassMAC as MAC protocol for the jammer node, because as we explained before the jamming attack interrupts the communication between the nodes, and makes them not able to access the channel, by perturbing the MAC protocol mechanism.

4.2.1. Simulation results and discussion

In this section, we measure the performance of our proposed approach jamming detection in WBAN, by calculating the network parameters in both cases without and with jammer node.

As shown in Figure 5 and Figure 6, when the legitimate node is under Jamming, the PDR value is very lower than the normal case, and the signal indicator is higher (-63,43 dBm). Moreover, in the normal case, the PDR parameter is higher and the RSSI is lower (-90 dBm). When the node (1) is near to the node (2) who is under jamming, we found that the node (1) could be also affected by the attack.

When the medical sensor (node 1) is under jamming, it consumes a lot of energy comparing with normal case (ECAth) as shown in the Figure 7.

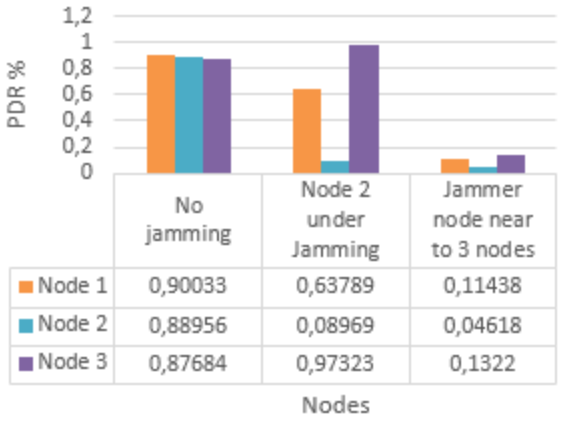


Figure 5: Packet delivery ratio values per node in WBAN (with and without jamming)

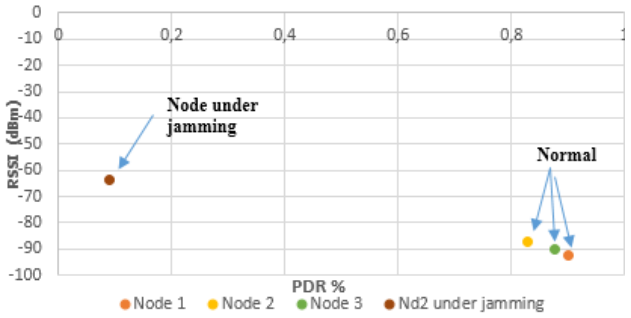


Figure 6: RSSI and PDR values in WBAN

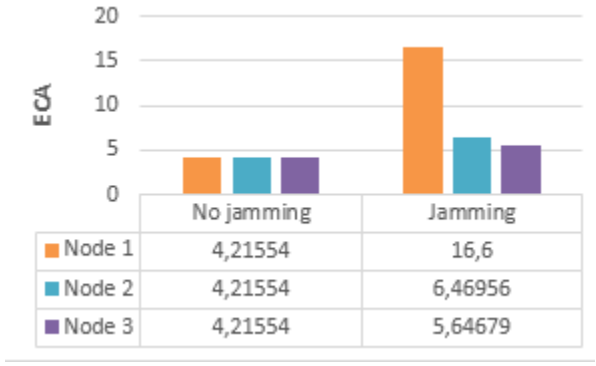


Figure 7: ECA parameter values in WBAN

Therefore, we can notice from these results that the used network parameters PDR, ECA, BPR and RSSI are very useful and have a great effect for identifying the presence of jamming attack in the WBAN system.

5. Conclusion and future work

In the recent years, many studies are interested to ameliorate and secure the healthcare industry. This paper is divided into two main parts. The first one presents our new classification of attacks in WBAN. The second one concerns a new IDS to detect three types of jamming attacks.

Thanks to our classification based on various measures, the researchers could have a deep view that helps them to understand the different attacks in order to propose a robust solution security for defending attacks in the WBAN system.

Our novel Jamming Intrusion Detection System was implemented on the medical sensors, which is based on four fundamental network parameters: PDR, RSSI, ECA and BPR. Indeed, by using a simple technique, based on these parameters, the jamming attack could identify three types of jamming as constant, deceptive and reactive. Furthermore, our mechanism detects these attacks without consuming a lot of energy. However, according to the best of our knowledge, the most of proposed techniques have some limits as consuming a lot of energy because of complex calculation techniques, detecting false alerts due to a less of parameters, and other IDS do not specify types of jamming. Our simulation carried out in two main scenarios, the first one is with jamming and the second one is without jamming, which is done with the popular WBAN's simulator tools OMNET++ and Castalia. According to the results of simulation, our method proves that the used parameters have good effects to detect the presence of jamming with less false alerts.

For future works, we aim to improve our proposed IDS by using Fuzzy Logic to make our mechanism more robust in order to increase detection and decrease the false alerts.

Acknowledgment

This research is performed inside the MATSI Lab., ESTO, University Mohammed First, Oujda (Morocco).

References

[1] Y. Qu, G. Zheng, H. Ma, X. Wang, B. Ji, and H. Wu, "A survey of routing protocols in WBAN for healthcare applications," *Sensors (Switzerland)*, vol. 19, no. 7, 2019.

[2] H. Fouad, "Continuous Health-monitoring for early Detection of Patient by Web Telemedicine System," pp. 76–83, 2001.

[3] S. Movassaghi, A. Mehran, J. Lipman, D. Smith, and A. Jamalipour, "Wireless Body Area Networks: A Survey," *IEEE Commun. Surv. TUTORIALS*, vol. 16, pp. 1658–1686, 2014.

[4] H. I. Reyes and N. Kaabouch, "Jamming and Lost Link Detection in Wireless Networks with Fuzzy Logic," *Int. J. Sci. Eng. Res.*, vol. 4, no. 2, pp. 1–7, 2013.

[5] N. A. Noureldien, "A novel taxonomy of MANET attacks," *Proc. 2015 Int. Conf. Electr. Inf. Technol. ICEIT 2015*, pp. 109–113, 2015.

[6] Z. Wu, Y. Ou, and Y. Liu, "A taxonomy of network and computer attacks based on responses," *Proc. - 2011 Int. Conf. Inf. Technol. Comput. Eng. Manag. Sci. ICM 2011*, vol. 1, pp. 26–29, 2011.

[7] Y. M. Amin, A. T. Abdel-hamid, and S. Member, "Classification and analysis of IEEE 802.15.4 PHY layer attacks," *2016 Int. Conf. Sel. Top. Mob. Wirel. Netw.*, pp. 74–79, 2016.

[8] Y. M. Amin, A. T. Abdel-hamid, and S. Member, "Classification and Analysis of IEEE 802.15.4 MAC Layer Attacks," pp. 74–79, 2016.

[9] P. Chahal, "Comparative Analysis of Various Attacks on MANET," vol. 111, no. 12, pp. 42–46, 2015.

[10] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The feasibility of launching and detecting jamming attacks in wireless networks," p. 46, 2005.

[11] B. Yu and L. Y. Zhang, "An improved detection method for different types of jamming attacks in wireless networks," *2014 2nd Int. Conf. Syst. Informatics, ICSAI 2014*, no. Icsai, pp. 553–558, 2015.

[12] V. C. Manju and K. M. Sasi, "Detection of jamming style DoS attack in Wireless Sensor Network," *Proc. 2012 2nd IEEE Int. Conf. Parallel, Distrib. Grid Comput. PDGC 2012*, pp. 563–567, 2012.

[13] M. Messai, "Classification of Attacks in Wireless Sensor Networks," *Icta*, no. April 2014, pp. 23–24, 2014.

[14] C. Del-Valle-Soto, L. J. Valdivia, and J. C. Rosas-Caro, "Novel detection methods for securing wireless sensor network performance under intrusion jamming," *CONIELECOMP 2019 - 2019 Int. Conf. Electron. Commun. Comput.*, pp. 1–8, 2019.

[15] B. Kiruthika, S. Abinaya, R. Ezhilarasie, and A. Umamakeswari, "Security

attacks and its countermeasures in wireless sensor networks - A survey," *Res. J. Pharm. Biol. Chem. Sci.*, vol. 6, no. 3, pp. 1374–1387, 2015.

[16] M. Çakiroğlu and A. T. Öozcerit, "Design and evaluation of a query-based jamming detection algorithm for wireless sensor networks," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 19, no. 1, pp. 1–19, 2011.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 86-92 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Modelling the Methodology to Assess the Effectiveness of Distributed Information Systems

Ivan Aleksandrovich Tarkhanov[1,2,*], Galina Pavlovna Akimova[1], Matvey Aleksandrovich Pashkin[1], Aleksander Vladimirovich Soloviev[1]

[1]*Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, 117317, Moscow, Russia*

[2]*State Academic University for Humanities, 119049, Moscow, Russia*

ARTICLE INFO

Article history:
Received: 21 September, 2019
Accepted: 23 December, 2019
Online: 15 January, 2020

Keywords:
Effectiveness
Reliability
Information system
Geographical distribution
Reliability scheme
Prim algorithm

ABSTRACT

This research aims to develop the methodology to assess effectiveness of distributed information systems (hereinafter IS). The effectiveness of information systems is reviewed along with their reliability. The article underlines the importance of reliability assessment for large geographically distributed systems. We suggest the approach based on the hierarchical presentation of reliability scheme that constitutes the core of the mathematical model for reliability of a distributed system. The key advantage of the suggested approach is its applicability on geographically distributed systems of different architecture. Successful testing during effectiveness and reliability assessment on 3 geographically distributed IS allows to suggest that the developed methodology is ready to be used in IS reliability assessment projects.

1. Introduction

Rapid development of information and communication technologies contributed to the design and development of new large-scale geographically distributed systems that have become an integral part in the progress of state and transnational corporations. A large-scale geographically distributed system can be viewed as a set of interconnected hard and software. If an IS is unreliable or ineffective then a failure of one or several of the critical nodes can crash the whole system. Thus, it is crucial to be able to detect risks in due time and check operation performance of an IS. Regular reliability and effectiveness assessment will help to eliminate these risks [1]. In large geographically distributed IS, these problems are aggravated by the complexity and heterogeneity of IS hardware and software architecture. Thus, effectiveness and reliability assessment becomes especially important [2].

However, if the term reliability is quite clearly defined [3] then the term IS effectiveness is rather vague. Thus, according to the Russian state standard - GOST 34.003–90 – the effectiveness of an automated system means «the extent to which the goal set at the development of an IS is reached» [4]. In order to be able to measure IS effectiveness at the start of IS design developers have to define quality parameters by which system effectiveness will be assessed during operation [5], [6].

The following parameters are specified when an IS operational effectiveness is discussed from a technical point of view [7]:

- Adaptability – the IS property defining the ability to correlate to different functionalities and change accordingly;
- Availability - the IS property defining the availability of system functions and data stored in it at any time;
- Response time - the IS property defining the ability to guarantee timeliness and appropriate performance with consideration of declared performance indicators [8];
- IS reliability;
- Some studies identify other parameters, e.g. usability [9] or information security [10], resistance to external actions, including disaster recovery etc.

Practice has proved that the system of the abovementioned parameters is quite universal [7]. Besides, if we set one of the parameters as a primary one, e.g. reliability, then the rest of the parameters will be considered as restrictions. Alternatively, reliability may be considered as a restriction to other quality parameters. For example, the system may be "absolutely reliable"

* Ivan Aleksandrovich Tarkhanov, Email: itarkhanov@gaugn.ru

but it does not ensure availability of data or timeliness (reasonable response time), thus it cannot be considered effective.

Over the last 25 years there have been many researches dedicated to reliability in the context of IS architecture, data inside this IS and automated processes [5], [6]. However, none of these researches allows to clearly correlate the system reliability with certain distributed nodes and get a unified assessment that can be used in different projects. This article continues the reliability assessment research [11] and is dedicated to the modelling of methodology to assess effectiveness of geographically distributed information systems. The main quality parameter for the purposes of this article is IS reliability, while other parameters are regarded as restrictions.

2. Modelling the reliability of geographically distributed information systems

Reliability is definitely an important parameter of IS effectiveness. According to [12], reliability is an IS property enabling to preserve in time assigned parameters required to fulfill the main function under the influence of malfunctions (failures and breakdowns) of hardware, software and data errors, personnel and users mistakes in standard operational conditions and working environment when maintenance and servicing parameters are specified.

IS consisting of many parts distributed over a large territory have much longer operational lifespan than general life cycle for computer technologies that is normally 5-7 years [13, 14]. For example, reviewed herein IS - Russian Federation Automated System "Vybory" and "Electronic archive for personal record keeping" of the Russian State Pension Fund –have been operating not less than 20 years [15].

2.1. Task setting for modelling of the methodology to assess effectiveness of distributed information systems

IS reliability assessment in a dynamic environment with the purpose of detection of vulnerable spots in terms of reliability will allow to identify beforehand the areas of possible failures. Besides, it will enable to make a targeted modernization of an IS, and, finally, to save money. Providing this, the task setting for IS reliability assessment method can be formulated as follows:

Given:

1. The set of soft and hardware of an IS (TSSH) $PT = \{ PT_i \}$
2. Parametric perturbations on the PT set in time t:
3. $\varepsilon(t)$ – reliability violation on the PT set,
4. Requirements towards reliability parameters $T = \{ \rho T \}$ on the PT set

Find:

- A set of reliability parameters assessment models $PT : \rho$

Solution:

- Make an IS reliability scheme.
- Develop a general reliability model for this scheme.
- Develop a reliability assessment method.

Supplementary notes:

Reliability can serve as a restriction for other IS quality parameters: availability, response time etc.

Within this task, we do not consider the issues of mass nodes failures of a distributed IS for reasons of catastrophic character.

2.2. The modelling of a reliability scheme for geographically distributed information systems

The development of reliability scheme is the key stage at effectiveness modelling in general and reliability modelling specifically. It is convenient to introduce a reliability scheme for hierarchical geographically distributed IS in the form of graph $G(V, E)$, where $V = \{ v_i \}$ is a set of reliability nodes v_i. Every node consists of all soft and hardware stored on a certain closed territory (an office, a building, an organization). $E = \{ e_j \}$ is a set of communication channels (CC) e_j. Every element of the set E connects a couple of elements of the set V.

The development of a mathematical reliability assessment model by (network) graph is a difficult task. However, the majority of geographically distributed IS is designed hierarchically, which makes the task considerably easier.

Statement. Reliability scheme of a large hierarchical geographically distributed IS can be introduced in the form of a tree or a forest.

Let us consider an IS structure consisting of one central node and two regional nodes. Let us assume that CC e_j have communication bandwidth of 4 Mb/s on the region- center level, 2 Mb/s between regions. The formed reliability graph is a network consisting of three nodes and three arcs (see fig.1). Let us assume that data processing power on v_i nodes of the center-region graph correlates as 2:1 (proportionally to the number of lower-level nodes to one upper-level node). It is useless to make evenly powerful nodes because this will result in overspending at the development of an IS.

The use of Prim algorithm [15] to form a graph tree (in our case a maximum tree) with the starting vertex in the root of the tree (vertex Center) produces a spanning tree without region-region arc. Maximum tree is formed because there is a need to create communication channels that will guarantee implementation of timeliness restrictions at reliability assessment.

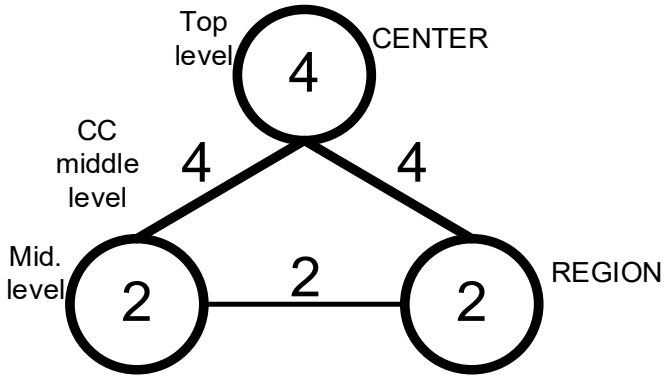


Figure 1. Reliability graph for basic two-level IS scheme

Increasing the number of levels in the IS hierarchy by one will produce an IS containing 3 levels of reliability nodes hierarchy v_i: center, region, district. In this IS every region has two districts. Let us assume that CC e_j have communication bandwidth of 4 Mb/s on the region-center level, 2 Mb/s on the district – region level, 1 Mb/s between districts of the same region, and 2 Mb/s between regions.

The formed reliability graph is a network consisting of seven nodes and nine arcs (see fig.2). Let us assume that data processing power on v_i nodes of the center-region-district graph correlates as

4:2:1 (proportionally to the number of lower-level nodes to one upper-level node).

Using the same modified Prim algorithm [16] we produce a spanning graph tree without arcs region-region and district-district. Thus, we got the same result for a three-level scheme as for a two-level one (providing we take into account the restriction that CC between nodes of the same hierarchy level are less powerful than CC between nodes of different hierarchy levels).

By applying a mathematical induction method, i.e. increasing the number of levels by addition of child nodes to the leaf node (fig.2) it is possible to form more complex graphs and maximum trees for them.

Providing the above we can formulate the task for reliability scheme modelling in the form of a tree or a forest based on the IS graph scheme.

Given:

A graph describing the IS scheme $G(V, E)$. Where $V = \{ v_i \}$ is a set of IS reliability nodes . Every node $v_i = \{ TSSH_{ik} \}$, $TSSH_{ik}$ is a certain set of soft and hardware of a node. $E = \{ e_j \}$ is a weighted set of all communication channels (CC) e_j. Every element of the set E connects a couple of elements of the set V. The weight of e_j means CC communication bandwidth.

Find:

The set of tree-type reliability schemes $Tr = \{ G'_l(V, E') \}$, where $G'_l(V, E')$ is the maximum tree l, $V = \{ v_i \}$ is the set of all reliability nodes v_i. $E' = \{ e_j \}$ is the set of IS communication channels (CC) of the maximum total weight that does not form a cycle. Every element of the set E' connects a couple of elements of the set V.

Restrictions:

- IS is hierarchically designed;
- The tree is formed by passing the arcs $G(V, E)$ from the root to the leaves (downwards) with the ban to pass the arcs upwards along the tree.

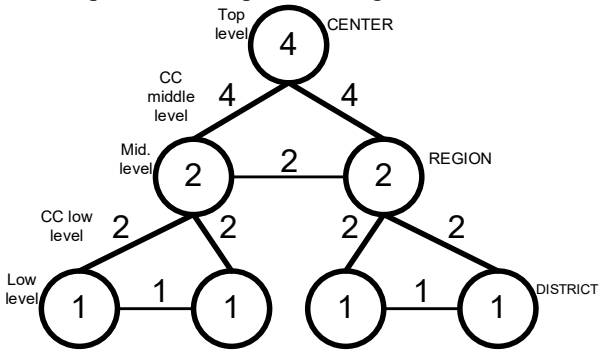


Figure 2. Reliability graph for a basic three-level IS scheme.

When a graph has a complex structure (see fig. 3), it may be difficult to locate a maximum tree (trees) because the use of classic Prim algorithm allows inclusion into the tree the arcs from lower levels to upper levels. This is possible, for example, when a starting vertex is Center 3 and there is no CC 5 (see fig.3). It is also possible to get a tree without arcs from vertex Center 3 to regions (see fig.3). Such solution will contradict the flow of main data in an IS.

For such cases, the Prim algorithm should be modified either by addition of the option allowing marking vertices in the data structure as non-visitable (e.g. in case of a failure) or the introduction into the algorithm of a vertex level label and a forced ban to consider arcs from lower to upper levels of the hierarchy (see fig.4).

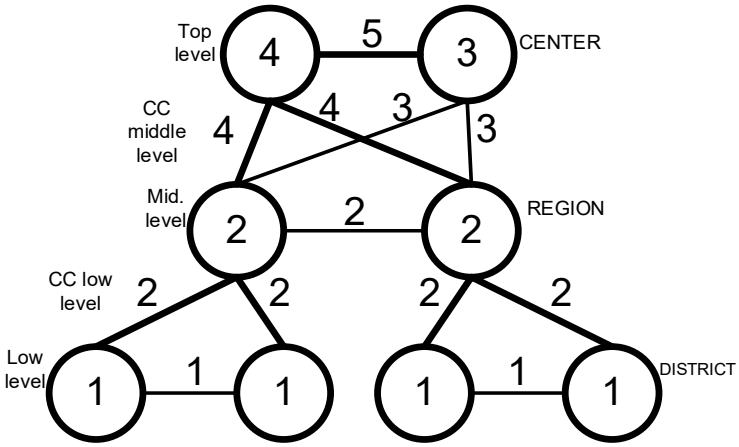


Figure 3. Reliability graph for a three-level IS scheme with 2 centers

If there is a backup data processing center for the central node (e.g. nodes Center 4 and 3 on fig. 3) then end-user processes are still addressing the same data copy at a time. Thus, the central node is considered as single (e.g. by combining vertexes Center 4 and 3 on fig. 3) but in the model it is necessary to consider its backup (e.g. in the basic case using the formula $K_{r.b.} = 1 - (1 - K_r)^n$, where n is the number of backup nodes for which reliability is defined as K_r).

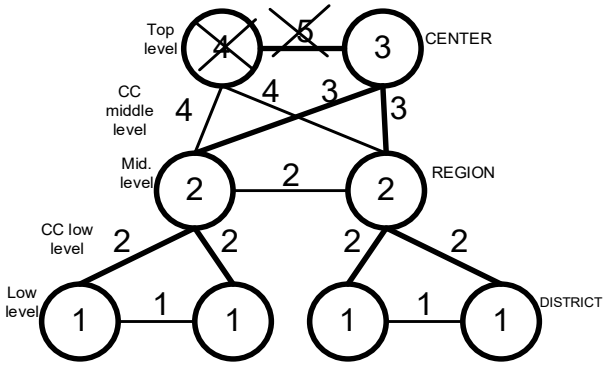


Figure 4. Reliability graph for a three-level IS scheme with two centers and the option to eliminate some vertices and arcs

If central nodes and communication channels have nominally the same processing power, it is possible to form several different trees if we choose different starting graph vertices, i.e. it is possible to form not a tree but a forest. Forming a forest does not significantly complicate the initial task. According to the reliability assessment model presented below it is necessary to make calculations by several trees and take the minimal assessment.

Forming one or several maximum trees (critical sections) in the general reliability graph is justified because the calculation and reliability scheme is much easier and there is no significant distortion of reliability assessment parameters.

Generally, for every tree (or section) of the initial reliability scheme there is a need to assess reliability according to the further suggested reliability assessment model and choose a minimal model. The complexity of reliability scheme algorithmic modelling depends on the number of vertices in every maximum

tree and in general case is $O(mn^2)$, where n is the number of vertices in a tree, and m is a number of trees.

The authors suggested this approach after the analysis of application of reliability assessment methods for large geographically distributed IS [17].

2.3. General reliability assessment model

IS reliability model for v_i nodes in general is formed based on the reliability scheme. Then, it is detailed for every v_i node depending on available soft and hardware $PT_i = \{ PT_{ij} \}$ of the v_i node.

The general reliability model scheme for a three-level hierarchical distributed IS (see fig.2) is as follows [18]:

$$\rho = K_{r\,IS} = K_{r\,t.l.} \sum_{i=1}^{Nm.l.} (b_i K_{r\,cc\,m.l.\,i} K_{r\,m.l.\,i} (\sum_{j=1}^{Nl.l.i} a_{ij} K_{r\,cc\,l.l.\,ij} K_{r\,l.l.\,ij}))$$

where $K_{r\,t.l.}$ – a reliability parameter for IS top-level objects (a non-detailed object reliability model for objects like data processing centers, central servers, software of central database etc.);

$b_i = Ob_i / \sum_{i=1}^{Nm.l.} Ob_i$ – the share of objects serviced by i- node of the middle level IS, Ob_i – a number of reliability elements serviced by i-node of the middle level IS ($\sum_{i=1}^{Nm.l.} b_i = 1$);

$N_{m.l.}$ – the number of elements of the middle level IS;

$K_{r\,cc\,m.l.\,i}$ – TSSH CC reliability parameter between top and middle level;

$K_{r\,m.l.\,i}$ –parameter of i-node reliability of the middle level IS;

$a_{ij} = Ob_{ij} / \sum_{j=1}^{Nl.l.i} Ob_{ij}$ – the share of objects serviced by j-node of the low level IS node connected with the middle level i ($\sum_{j=1}^{N\,l.l.i} a_{ij} = 1$) against general number of low level objects connected with i-node of the middle level IS;

$N_{l.l.i}$ – the number of low-level nodes connected with i-node of the middle level;

$K_{r\,cc\,l.l.\,ij}$ - the TSSH CC reliability parameter of the middle level i-node – low level j-node;

$K_{r\,l.l.\,ij}$ – the reliability parameter of the low level j-node connected with the middle level i-node.

A general model for reliability assessment is formed for availability parameter K_r (probability that the IS will be functioning at any given time), i.e. for a complex parameter for failure-free performance and recoverability of an IS. However, the choice of a reliability parameter is not crucial in this situation – the methodology shows how to design a reliability scheme, how to transform it into a tree (or a forest), how to form a reliability assessment model according to the scheme (schemes). Similarly, it is possible to form a reliability models for other parameters, such as:

- Operational availability parameter (K_{or}) is the probability that the object will be functioning at any given time t, apart from planned periods when the use of the object for its intended purpose is not required. Starting that moment the object will demonstrate failure-free performance for the set period.
- Technical use parameter (K_{tu}) is the ratio of expected value of intervals when the object is operationally available over a period of use to the sum of expected values of intervals when the object is available and idle due to technical maintenance or repairs over the same period of use.
- Planned use parameter (K_n) is a part of use period when the object must not be on planned maintenance or repair.
- Effectiveness parameter (K_{ef}) is a ratio of the value of effectiveness parameter when the object is used for its intended purposes for a certain period of operation to the nominal value of this parameter calculated under the condition that the object failures does not happen over the same period. The effectiveness parameter specifies the influence of object failures on its effectiveness when used for its intended purposes.

The suggested model for reliability assessment allows by induction to increase or decrease the number of levels, by calculation of corresponding weighting factors and by addition of the next level standard reliability parameters sum. It should be noted that the failure of a node of any level equals to the failure of all its child nodes. This partly takes into account mass failures.

3. Reliability assessment methodology

Now that the mathematical reliability assessment model is developed, it is time to start modelling the methodology to ensure reliability. It should be started with the elaboration of the terms failure and crash for a certain system, the choice of reliability parameters (availability rate, probability of failure-free performance, operational availability rate).

For the purpose of this research, IS is considered as a complex geographically distributed multifunctional system. Thus, in addition to «Full availability» and «Total crash» states there are many intermediate states with different levels of availability and corresponding performance effectiveness.

The developed model allows defining an IS total crash state - a total system crash is possible at the failure of upper level object and/or failure of all middle level objects and/or all low level objects.

Partial IS failure – the failure of one middle level object, which equals to the failure of all lower level objects handled by this middle level sector of an IS.

3.1. Modelling the methodology for reliability assessment

Summing up the methodology modelling in three major projects (Russian State Automated System "Vybory" [15], Electronic archive for personal record keeping" of the Russian Pension Fund and management system for the concept «Smart City Skolkovo» [16]) we came to the conclusion that it is necessary to undertake the following steps to develop a quantitative reliability assessment methodology and make the actual assessment:

1) Develop a reliability scheme in the form of a graph and detect one or several trees.
2) For every formed tree develop a mathematical model for reliability assessment according to the presented above model specifying the number of node levels.
3) Elaborate the terms for failure and crash of an IS according to its functionality. For example, the system failure is registered if function $f_i \in F$ cannot be executed for a period of time more than T minutes (seconds), where F is a set of functions, crucial to IS if not executed.
4) Define the terms failure and crash according to the critical number of IS operating nodes. For example, the failure of an IS node is registered if less than N automated working

stations (AWS) of an IS does not perform required functionality F for a period of time more than T minutes (seconds).

5) Define reliability parameter (parameters), principles of failure and crash probability distributions, restrictions and assumptions at reliability assessment.
6) Define the effect other IS have on the system reliability and adjust the mathematical model taking into account this effect.
7) Develop the form and procedure for the collection of IS reliability statistics data, which include every node and element of IS scheme and reliability model.
8) Collect statistic data on the nodes and reliability elements of the IS.
9) It is recommended to develop mathematical models for reliability parameters, e.g. for timeliness and IS function execution, and the scheme for the use of these parameters at the system reliability assessment. For example, if at the current reliability rate timeliness is not sufficient, then the failure of a node or IS in general is registered. Thus, in the most general case reliability parameters serve as reliability model restrictions.
10) Specify the general mathematical reliability model by every node and reliability element taking into account possibility of backup and influence of other parameters and systems.
11) Calculate system reliability parameters taking into account restrictions (including timeliness and function execution). The calculations are made for every formed i-tree. Then a pessimistic approach is applied, i.e. $K_r\ IS = min(K_r\ IS_i)$ is chosen.
12) Calculate reliability parameters according to the model described above for every node of every level of the reliability scheme.
13) For reliability nodes for which there are no failure statistics calculate anticipated reliability as the lowest among the same calculated reliability nodes of an IS.
14) If necessary, specify parameters calculations for every node or reliability element.
15) At the end of the calculation make an IS "health chart", i.e. IS scheme where risky nodes and reliability elements are marked.
16) The overall risk and weaknesses chart used at the detailed reliability assessment is defined by detected critical nodes.
17) If the risks are confirmed by the detailed assessment, it is necessary to develop an action plan to eliminate these risks.

3.2. Specifics of the interpretation of calculation results in the reliability model

Consider a more complex case. A distributed IS consists of several top level nodes, thus it is necessary to form a maximum tree for every vertex and calculate the assessment for every such tree.

As a result we have a set of assessments: $K_{r\ IS} = \{ K_{r\ ISi} \}$.

In IS reliability assessment we apply a pessimistic approach, i.e. choose $K_{r\ IS} = \min_i(K_{r\ ISi})$.

The produced indicators both for IS in general and for every separate reliability node and element are compared against the reliability set $T = \{ \rho_T \}$ existing on the PT set.

Reliability assessment can be done quite fast with the use of the suggested method. And with the use of some software for calculations even faster. However, the assessment is not too accurate and can be used only for estimated rough calculations.

4. Modelling the effectiveness of distributed information systems

In order to assess effectiveness it is needed not only to model reliability but also develop models of other IS quality performance parameters.

The authors developed the following mathematical models of IS quality parameters for the assessment of IS effectiveness: IS availability and response time (timeliness). These models can be supplemented by IS reliability assessment models as additional restrictions.

4.1. The model for the assessment of IS functions performance availability

Here is the formal task formulation for the IS availability assessment by listed functions.

Given:

1) Functionality volume F (total number of all IS operations by all operation types).
2) Data volume D (total data volume stored in IS database).
3) Restriction F_d (maximum admissible F value that allows to consider IS functional).
4) Restriction D_d (maximum admissible D value that allows to consider IS functional).
5) IS reliability (K_r).

Find:

Probability of availability of the minimum admissible IS functionality and data volume.

Solution:

Mathematical model of availability is the probability to perform IS operation s by a user:

$$P_p\,(F \geq F_d) = P(F \geq F_d)\ K_{r\ IS}$$

And the probability of IS data availability to a user:

$$P_d\,(D \geq D_d) = P(D \geq D_d)\ K_{r\ IS}$$

where $P(F \geq F_d)$ – the probability of performing functionality volume F, i.e. performing of IS functions not less than directly set (or admissible) F_d subject to failure-free performance of IS soft and hardware;

$P(D \geq D_d)$ – the probability of availability of data volume D not less than directly set (or admissible) D_d subject to failure-free performance of IS soft and hardware;

$K_{r\ IS}$ – IS reliability indicator calculated according to the reliability assessment model (see above the general IS reliability assessment) that serves as a restriction of the model.

$$P(F \geq F_d) = n_i / N(i),$$

where n_i – the number of operation of i-type (see the list of operation types above) that satisfies the inequation $F \geq F_d$,

$N(i)$ – the total number of operation of *i-type* performed in an IS at the collection and processing of information.

$$P(D \geq D_d) = d / D,$$

where d – the volume of available data,

D – total data volume in an IS.

Additional parameters specifying availability:

- Average percent of IS functions F availability over a period of T;
- Average percent of IS data D availability over a period T.

4.2. The model for the assessment of IS functions timeliness (response time)

Here is the formal task formulation for the IS timeliness assessment.

Given:

1) Functionality volume F (total number of all IS operations by all operation types).
2) Restriction T_d (maximum admissible time value that provides an IS user with acceptable result of the IS function performance)
3) IS reliability ($K_{r\,IS}$)

Find:

The probability of IS performance of its functionality in full over a period not less than maximum admissible time

Solution:

Mathematical model for timeliness is defined as the probability of information processing over a period of t that is not more than the set (admissible) T_d, with reliability restriction:

$$P_t\ (t \le T_d) = P\ (t \le T_d)\ K_{r\,IS}\ ,$$

where $P\ (t \le T_d)$ – the probability of information processing during operations performance (the types of operations are listed in the first paragraph of the section) in an IS over a period of t that is mot more than the set (admissible) T_d subject to failure-free performance of soft and hardware;

$K_{r\,IS}$ – IS availability indicator (see section 2.4).

In order to calculate all time intervals T it is necessary to obtain statistics data and find admissible time T_d allowing to get the probability assessment of function timeliness.

Probability assessment of i-type function timeliness over a period not exceeding admissible is calculated by the following formula:

$$P(T_i \le T_{di}) = n_i / N(i)$$

where n_i – the number of performed i-type operations T_i that satisfy the inequation $T_i \le T_{di}$;

$N(i)$ – the number of operations of i-type from the set of operations $F\,l = [1, N(i)]$.

5. Practical implementation of reliability assessment methodology

Effectiveness and reliability assessment methods developed according to the suggested by the authors methodology are described in [13, 18]. This methodology can be easily automated. For example, software complex for the assessment of effectiveness and reliability of the Russian Federation Automated System "Vybory" was created based on this methodology.

The suggested methodology formed the basis for the methods used by the Russian Federation Automated System "Vybory" and soft and hardware complex "Electronic archive for personal record keeping" of the Russian Federation Pension Fund. The calculations made according to the implemented methods revealed reliability bottlenecks in IS nodes, which allowed to make their targeted modernization.

Besides, the developed methodology was used at the development of the logic-mathematical model for Skolkovo "Smart city" management system concept for the Skolkovo Innovation center [17]. This allowed to design the management system with regard to reliability and effectiveness on the concept stage and set justified criteria for the selection of soft and hardware solutions at the detailed design stage.

6. Comparison with other approaches and model restrictions

First researches on the reliability assessment of distributed systems were focused on the solution of individual tasks - assessment of mistakes at data transfer or instability of communication channels bandwidth between IS nodes [19],[20]. Engineering approach and special set of tools to assess IS reliability is described in [5] и [6].

The most detailed IS effectiveness assessment model that takes into account not only the quality of the system itself but also the data circulating in it, user experience and other business-oriented parameters is D&M IS Success Model [7]. According to classification [7] the described model assess only technical aspect of the system performance and is similar to approaches based on IS architecture assessment [20], [21]. It can be used along with D&M IS Success Model and supplement it with regard to geographically distributed systems.

Another possible application of the suggested method is its use with IS effectiveness assessment models based on machine learning and data analyses [22]. During machine learning the reliability scheme in force (or any other parameter of IS effectiveness) is adjusted and used for the presentation of the current situation and fast calculation of IS effectiveness parameters.

The majority of researches focuses on "reliability of data" circulating in the system (a process approach) [23] or on failure probability [24] not taking into account architecture specifics of the assessed IS.

The most recent researches on the reliability assessment of distributed systems are mainly based on the use of Markov chains [2], [25] that allow to understand the correlation between the states of different system components, define bottlenecks of a small distributed system. However, it does not allow to get a clear IS reliability assessment comprising over 100 nodes in a dynamic state.

In section 1.2 it is demonstrated that reliability scheme modelling with the use of Prim algorithm is applied for hierarchical systems. The same approach may be employed in the non-hierarchical systems. To do so it is needed to locate one or several central nodes and form the reliability scheme for each of them.

It should be noted that the suggested model does not take into account some important parameters, e.g. it does not cover the issue of data deterioration at the reliability assessment [26] that in future could be considered at the development of reliability scheme or any other parameter of IS effectiveness. A simplified description of the mathematical model was presented in the suggested methodology.

7. Conclusion

This article describes methodology for modelling IS effectiveness that was brought up to the level of applied operational methods that were implemented during the development of the major federal level information systems: the Russian Federation State Automated System "Vybory", "Electronic archive for personal record keeping" of the Russian Federation Pension Fund, Skolkovo "Smart city" concept, which is undoubtedly the advantage of this work.

Another advantage of the developed mathematical model is that it takes into consideration a geographically distributed architecture of an IS and has a great potential for extensibility into more complex distributed IS configuration.

In further studies, it is planned to improve Prim algorithm convergence at the formation of critical sections (maximum trees of the reliability scheme), to develop a method for specification of nodes reliability models of the reliability scheme and its automatization with the help of specialized software, to develop a methodology to model other IS quality parameters such as adaptability, resistance to external actions (including actions of catastrophic character).

We plan to develop the software allowing to improve the reliability scheme calculation of certain nodes and elaborate modelling methodology for such effectiveness parameters as adaptability and disaster recovery. The focus of the research will be on the analysis of Prim algorithm convergence at the calculation of several maximum trees.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] L. F. Pitt, R. T. Watson, C. B. Kavan "Service quality: a measure of information systems effectiveness" MIS quarterly, 173-187, 1995. https://doi.org/10.2307/249687

[2] A. Waseem, Y.W. Wu "A survey on reliability in distributed systems" Journal of Computer and System Sciences, No. 79.8., 1243-1255, 2013

[3] F. M. Zahedi "Reliability metric for information systems based on customer requirements" International Journal of Quality & Reliability Management 14.8, P. 791-813, 1997. https://doi.org/10.1108/02656719710181312

[4] GOST 34.003-90 Informacionnaja tehnologija. Kompleks standartov na avtomatizirovannye sistemy. Avtomatizirovannye sistemy. Terminy i opredelenija, M.: Izd-vo standartov., 1992.

[5] M. C. Boudreau, D. Gefen, D.W. Straub "Validation in information systems research: a state-of-the-art assessment" MIS quarterly, 1-16, 2001. https://doi.org/10.2307/3250956

[6] P.K. Kapoor, H. Pham, A. Gupta, P.C. Jha "Software reliability assessment with OR applications" Springer, London, 2011.

[7] W. H. Delone, E. R. McLean "The DeLone and McLean model of information systems success: a ten-year update" Journal of management information systems. No. 19(4). 9-30, 2003. https://doi.org/10.1080/07421222.2003.11045748

[8] X. Lu, K. Moriyama, I Luque., M. Kanda, Y. Jiang, R.Takanuki, K. Mori "Timeliness and reliability oriented autonomous network-based information services integration in multi-agent systems" IEICE TRANSACTIONS on Information and Systems. No. 88(9), 2089-2097, 2005. https://doi.org/10.1093/ietisy/e88-d.9.2089

[9] A. Bangor, P. T. Kortum, J. T. Miller "An empirical evaluation of the system usability scale" Intl. Journal of Human–Computer Interaction. No. 24 (6), 574-594, 2008. https://doi.org/10.1080/10447310802205776

[10] M. Nancy, T. Stehney "Security quality requirements engineering (SQUARE) methodology" ACM SIGSOFT Software Engineering Notes. Vol. 30. No. 4. 2005. https://doi.org/10.1145/1082983.1083214

[11] Akimova G. P., Solovyev A. V., Tarkhanov I. A., Reliability assessment method for geographically distributed information systems // 2018 IEEE 12th International Conference on Application of Information and Communication Technologies (AICT), Almaty, Kazakhstan, 2018, pp. 1-4., https://doi.org/10.1109/ICAICT.2018.8747055

[12] GOST 27.002-89. "Reliability in technology. Basic concepts. Terms and Definitions", M.: Publishing Standards, 1989.

[13] P. Stafford "You can believe the Hype Cycle's take on technology" Financial Times. 30 May 2008.

[14] B. Furneaux, M. R. Wade "An exploration of organizational level information systems discontinuance intentions", Mis Quarterly, 573-598, 2011. https://doi.org/10.2307/23042797

[15] G. P. Akimova, E.V. Pashkina, A. V. Soloviev "Analiz ocenki jeffektivnosti ierarhicheskoj territorial'no-raspredelennoj informacionnoj sistemy na primere GAS "Vybory" Proceedings of the Institute of System Analysis of the Russian Academy of Sciences, vol. 58, 25-38, 2010.

[16] T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein "Introduction to Algorithms" Third Edition. MIT Press. ISBN 0-262-03384-4. Section 23.2: The algorithms of Kruskal and Prim, 631–663, 2009.

[17] A.V. Soloviev, Concept «Umnyj gorod Skolkovo» for Skolkovo innovation center and set of measures for its implementation. Vol 3. Logiko-matematicheskaja model' «Umnogo goroda»: research report. ISA RAN. 2012.

[18] A. V. Soloviev, E.Y. Komleva, "Metodicheskoe obespechenie nadezhnosti v oblasti hranenija jelektronnyh dokumentov" Proceedings of the XXIII International Conference "Documentation in the Information Society: Archival Studies and Documentation in the Modern World", Rosarchive, , 321-331, 2017.

[19] F. Proschan, R. E. Barlow "Mathematical theory of reliability." SIAM series in applied mathematics. 1967.

[20] L. Svobodova "Reliability issues in distributed information processing systems" Dig. Papers FTCS-9: 9th Annu. Symp. Fault-Tolerant Computing. 1979.

[21] K. Goševa-Popstojanova, S. T. Kishor "Architecture-based approach to reliability assessment of software systems" Performance Evaluation. No. 45.2-3. 179-204. 2001. https://doi.org/10.1016/s0166-5316(01)00034-7

[22] N. P. L. Ernawatiningsih, P. Kepramareni,, "Effectiveness of Accounting Information Systems and the Affecting Factors" International Journal of Applied Business and International Management, 4(2), 33-40, 2019 https://doi.org/10.32535/ijabim.v4i2.564

[23] R. Krishnan, J. Peters, R. Padman, D. Kaplan "On data reliability assessment in accounting information systems" Information Systems Research. No. 16(3). 307-326, 2005. https://doi.org/10.1287/isre.1050.0063

[24] A. Choudhary, A.S Baghel., O.P. Sangwan "Parameter Estimation of Software Reliability Model Using Firefly Optimization" Data Engineering and Intelligent Computing. Springer. Singapore, 407-415, 2018. https://doi.org/10.1007/978-981-10-3223-3_39

[25] Z. Zheng "Semi-markov models of composite web services for their performance, reliability and bottlenecks" IEEE Transactions on Services Computing, No. 10.3, 448-460, 2017. https://doi.org/10.1109/tsc.2015.2475957

[26] J. Xu, D. Yu, Q. Hu, M. Xie "A reliability assessment approach for systems with heterogeneous component information" Quality Engineering, 1-11, 18.12.2017. https://doi.org/10.1080/08982112.2017.1402935

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 93-99 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Harmonized European Drone Market? – New EU Rules on Unmanned Aircraft Systems

Anna Konert*, Tadeusz Dunin

[1]*Faculty of Law and Administration, Łazarski University, Warsaw 02-662, Poland*

ARTICLEINFO

Article history:
Received: 17 February , 2020
Accepted: 10 April, 2020
Online: 03 May, 2020

Keywords:
Air law
BVLOS operations
Drones
EU law
European Union drone market
Polish law
UAS
UAV
Unmanned aircraft systems
VLOS operations

ABSTRACT

The European drone market has been showing steady growth year after year. New EU drone rules will come into force as of July 1st, 2020, with the European Union setting itself the target to replace national rules with a common regulation with the ultimate goal of creating a harmonized European drone market. This study will clarify that the EU's regulatory framework covers all types of existing and future drone operations, creating an international market for unmanned aircraft services. This move will facilitate the enforcement of citizen's privacy rights and address security issues and environmental concerns to the benefit of the EU citizens. Moreover, this study will show that national legislators are now faced with the difficult task of replacing their national regulations with EU rules, however these were drafted so fast that they still leave a number of issues to be decided on by national legislators. The method of this study comprised of content analysis of existing legislation. The current doctrine were confronted with existing regulations, documents, and materials.

1. Introduction

The UAV and UAS markets have been showing steady, increasing growth year after year in Europe. Some of the world's top drone service providers are based here, and their results are proof of continuous and considerable market growth [1]. Some forecasts predict that the European drone market will more than double in size between the years 2018 and 2024 [2]. Others predict that revenues generated from commercial drone use in Europe will grow from USD 251 million in 2020 to USD 3 billion in 2025 [3] Different estimates see the region's drone market bringing in EUR 10 billion annually by 2035, and over EUR 15 billion annually by 2050 [4], with countries like Germany, France and the UK representing the lion's share of this total mainly due to their regulatory structure and the sheer size of their economies.

This predicted growth will most likely stem from an increase in popularity in commercial drone usage in many areas. According to SESAR (Single European Sky ATM Research), there are presently around 10,000 commercial UAV's in use in Europe – with forecasts for this number to rise to 200,000 in 2025, and up to 395,000 in 2035.

The non-military use of drones is already significant and extensive, spreading to areas never before though possible just a decade or so ago. UAVs are presently being used by the media (TV), in law enforcement and policing activities, in border patrols, for agricultural and environmental monitoring, by fire services, for power line surveying, aerial photography, and in security management systems (mostly in autonomous and semi-autonomous locomotion). For practical applications of UAVs in security management systems, please see "A Cyber-Vigilance System for Anti-Terrorist Drives Based on an Unmanned Aerial Vehicular Networking Signal Jammer for Specific Territorial Security" [5], "Design and implementation of a cyber-vigilance system for anti-terrorist drives based on an unmanned aerial vehicular networking signal jammer for specific territorial security" [6], "An Aerial Landmine Detection System with Dynamic Path and Explosion Mode Identification Features" [7]. Moreover, SESAR predicts that the application of UAVs will be more diverse as we move forward, with rising numbers of drones employed in the energy sector, public safety and security, mobility and transport, the media, insurance, e-commerce and delivery, real estate, academic research, telecommunications, and mining and construction [8].

*Corresponding Author: Anna Konert, a.konert@lazarski.edu.pl

It is no surprise that authorities across the Old Continent are scrambling to make Europe a powerhouse in terms of unmanned aerial business and a global first with regards to the mainstream and commercial use of UAVs. Take, for example, the U-Space initiative and its goal to implement a future air traffic management system that is open to and even capable of monitoring and directing drones (UAS Traffic Management, UTM [9]. Ambitious as this may sound, such initiatives get the ball rolling for a future that is integrated with unmanned vehicles.

As with every new technology, unmanned aircraft vehicles or systems possess their own regulatory and legal constraints. However, the main legal issue is of a regulatory nature, as the central aim of aviation regulations is to provide an adequate level of safety. In the case of UAVs, regulations aim to protect third parties on the ground and in the air, as there are no crew or passengers on board.

In order to enable the safe operation of remotely-controlled aircraft, or their operation by automated flight control systems, it is essential to develop detailed solutions for a number of technical and operational aspects.

This includes but is not limited to the following issues:

- airworthiness standards for UAV and UAS elements,
- data transmission standards for remote pilot stations, including protection against unauthorized interference etc.,
- collision avoidance systems, including motion and obstacle detection,
- emergency systems (e.g. loss of communications),
- unmanned traffic management systems,
- operator competences and training [10].

Drafting regulations that would enable the safe operation and development of unmanned aviation is no easy task. These need to be developed with a thorough understanding of the technology involved and with detailed knowledge of the risks that come with UAV operations. The scope of integration of UAVs with manned aviation is especially significant, as both operate within the same airspace. This is a long-term challenge for UAS, however it is co-dependent on improvements made in the field of manned aviation in the future [11].

From a regulatory point of view, the first step on the road to the integration of manned and unmanned flight is to allow for Visual Line of Sight Operations (VLOS). Most countries have already taken the appropriate steps for this by adopting or developing their national legislation to include drones, usually with weight limits up to 25 kg, in VLOS operations.

The next step forwards is for a regulator to allow Beyond Visual Line of Sight Operations (BVLOS) outside of segregated airspace or autonomous operations – which is now the case in Poland and some other EU countries.

When thinking about the European Union, it is all too convenient to think of its constituents as a single uniform marketplace. In reality, the EU it a group of drone technology markets each with their own regulations, and all of these have distinct specifications characteristic to their respective countries. Over the years, these regulations have been a barrier for the adoption of commercial drone technology across the region, with EASA (European Union Aviation Safety Agency) even claiming that they stifle business development and innovation [12]. But this is about to change with the agency's publication of a set of common, Pan-European regulations for UAV operations [13].

This paper will analyze current regulations in Poland and new EU regulations, showing how the region's drone market will change in the nearest future.

2. EU Regulations

"Europe will be the first region in the world to have a comprehensive set of rules ensuring safe, secure and sustainable operations of drones both, for commercial and leisure activities. Common rules will help foster investment, innovation and growth in this promising sector" – Patrick Ky, Executive Director of EASA [14].

In 2014, the European Commission published a communication entitled "A new era for aviation. Opening the aviation market to the civil use of remotely piloted aircraft systems in a safe and sustainable manner" [15].

Next, a legislative process aimed at extending the EU's competences in terms of safety regulations for drones and their operations was launched in 2015. In the opinion of EU institutions, the differences that exist between national regulations of Member States hamper market growth for the entire UAS sector.

By adopting Regulation (EC) No 216/2008 of the European Parliament and of the Council on common rules in the field of civil aviation, and by establishing the European Aviation Safety Agency [16] (through the so-called Basic Regulation on EASA), the EU set out to regulate matters related to the operation of civil unmanned aircraft with an operating mass of less than 150 kg. The establishment of regulations that concern unmanned aircraft used by the military, governmental services, average citizens, those designed for experimental purposes, and all other unmanned aircraft with an operating mass of no more than 150 kg fall within the competences of EU Member States. Such a state of affairs resulted from the fact that EASA's competences were extended in 2008 with the adoption of Regulation 216/2008, however the EU did not take into consideration the then unpredictable and rapid development of the unmanned aircraft sector [17].

On September 11th, 2018, Regulation (EC) No 216/2008 was repealed with the purpose of the new regulation establishing and maintaining a high and uniform level of civil aviation safety in all Member States, including the harmonization of regulations on the use of UAVs in all EU countries. Some of the main changes that were introduced included the obligation for operators to register their UAVs online (those weighing over 250 g) and common standards for UAV certification. However, these EU regulation serve as guidelines and do not contain specific provisions on the implementation of UAV operations, which were introduced by the Commission's delegated and implementing regulations.

As of July 2020, national rules in EU Member States will be replaced by a common EU regulation. Since neither the EU Parliament nor the EU Council had any objections, both Implementing and Delegated Acts (Commission Delegated Regulation (EU) 2019/945 [18] and Commission Implementing Regulation (EU) 2019/947 [19]) were published in mid-June of last year and entered into force 20 days later. These regulations will become gradually applicable one year from their date of publication, giving Member States and UAV operators time to prepare with the transitional period being fully completed by 2022. However, UAV operators will be obliged to register in the Member

State of their residence or there where they have registered their main place of business by as early as July 2020.

The new EU regulatory framework will cover all types of existing and future drone operations, enabling operators – once authorized in their state of registration – to freely circulate between Member States.

The purpose of introducing these new regulations is to ensure the safety of UAV operations, as well as protect the privacy of EU citizens, with respect to personal data protection, and the environment while allowing free access to airspace. The new regulations establish technical and operational requirements, provisions for UAS operations and personnel (minimum requirements and operator training), including both pilots and any organizations.

While maintaining its primary aim to ensure safe UAV operations, the new pan-EU regulatory framework will facilitate the enforcement of citizens' privacy rights and address security issues and environmental concerns. As touched on earlier, these new rules include both technical and operational requirements for UAVs and their operators. They define UAV capabilities, types of operations, and label these into three broad risk-based categories. These three categories of operations are based on the levels of risk involved per UAV flight and each adopts a varied regulatory approach, with UAV flight operational limitations decreasing with the requirement for greater authorization from a Member State's national aviation authority [20].

Regulation 2019/947 presents a comprehensive system of unified legal regulations which classifies UAVs into three categories based on the risks involved in their operations, their mass, and their application:

- Open. Operations in this category do not require official permission but are subject to a number of restrictions including those on UAV weight (up to 25 kg), types of operations (only VLOS, except when an unmanned aerial vehicle observer is involved), safety requirements (through the use of product safety regulations), obligations for operator registration, and geo-fencing systems. During Open operations, an operator must maintain a safe distance from other people and cannot fly over crowds, at a height of over 120 m, and cannot transport or drop hazardous materials. Open operations are further divided into three subcategories. (Article 4 of Regulation 2019/947);
- Specific. Operations in this category require authorization from aviation authorities based on a risk analysis of the proposed operations. This authorization can either concerns a single operation or a series of operations specified by time or place. Operation authorization must contain a relevant detailed list of risk mitigation measures. (Article 5 of Regulation 2019/947).
- Certified. Operations in this carry the highest risk. This category requires drone certification under Regulation (EU) 2019/945 as well as operator certification, which may, in some cases, include obtaining a drone pilot license. Certified operations include those performed under any of the following conditions: over crowds or gatherings, or involving the transport of persons or dangerous goods. In addition, the competent authority may, based on a risk assessment (Article 6 of Regulation 2019/947).

This new set of common rules (summarized in Table 1 and presented as a visual representation in Figure 1) will provide both professional and recreational drone operators with a much clearer understanding of what they are and are not allowed to do. Continent-wide and uniform regulations, especially the requirement that UAVs be registered and individually identifiable, will surely prevent events similar to the ones that took place in the UK at Gatwick and Heathrow airports back in 2018 – which grounded aircraft for up to 36 hours, affected over 140,000 passengers and cost airports an estimated GBP 50 million [21] – from occurring again.

Table 1: Authorization and regulations applicable to each of the adopted three categories of UAV operations [22]

Category of operations	Open *low risk*	Specific *medium risk*	Certified *high risk*
Authorisation needed	None	Authorisation from NAA based on operational risk assessment or specific scenario	Authorisation from NAA/EASA
UAS	Compliant with Commission Delegated Regulation on UAS	Compliant with requirements included in the authorisation	Certified UAS
Operations allowed	Restricted to: ▪ VLOS ▪ Altitude < 120 m ▪ Other limitations defined by: - Commission Regulation on UAS operations - National airspace zones	Restricted to: ▪ Operations specified in the authorisation ▪ Limitations defined by national airspace zones	Controlled airspace U-Space
Regulations	Commission Regulation on UAS operations in open and specific		Revision of existing aviation regulation
	Commission Delegated Regulation on UAS	No regulatory requirement (UAS requirements included in the authorisation)	

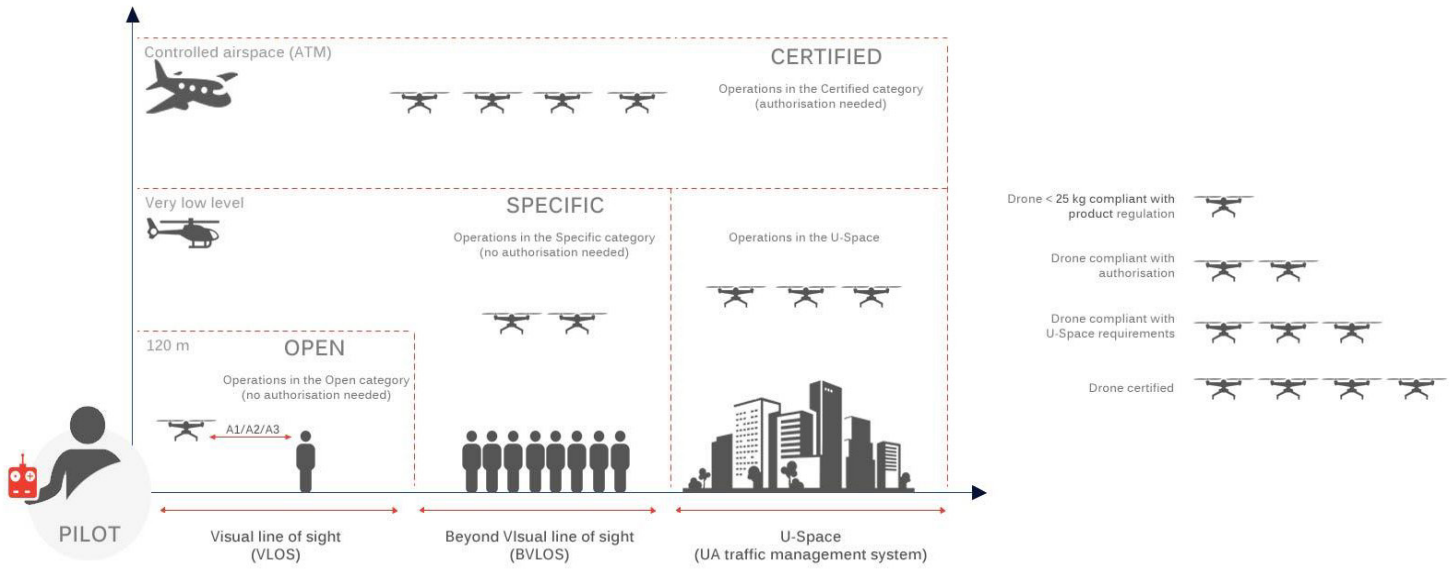
The benefits of facilitating uniform UAV operation requirements across all Member States seem clear. This is an unprecedented step forward, as even the United States are facing problems with UAV regulation unification due to the fact that individual states prefer to enforce local regulations [23].

Also, Member States will be able to define zones, most probably though satellite geo-location, where UAVs will not be permitted to enter or where they will be allowed more freedoms such as BVLOS flights [24], allowing countries to cater to their individual national specificities.

3. Discussion

This section starts with an analysis of Poland's national regulations. The first national regulations for UAVs were adopted in the country in July of 2011, in the Act of 3rd July 2002 – Aviation Law (Aviation Law Act) [25].

However, the very first implementing rules only concerned VLOS operations. The Regulation of the Minister of Transport, Construction and Maritime Economy of March 26th, 2013, on the exclusion of certain provisions of the Aviation Law Act – for certain types of aircraft and the specification of conditions and requirements for the use of these aircraft [27], was the Polish government's very first attempt to regulate general requirements for UAV operations, becoming the first national regulation for UAVs in Europe. This regulation was widely amended in 2016 [28], with the amendments being based on the practical aspects of unmanned operations.



A new regulation was adopted on December 20th, 2018 and entered into force on February 1st, 2019. This regulation amended previous rules from 2013 on the exclusion of certain provisions of the Aviation Law Act and included a new set of rules for BVLOS operations (MTOW of up to 25 kg, a maximum height of 120 m, and BVLOS operations outside of segregated airspace) [29].

This created new possibilities for the country's UAV sector, but also led to a series of safety challenges for manned aviation. It is still too early to make an assessment of the new regulation, as the first three months in which it has been in force saw the performance of just 20 UAV BVLOS flights under the new rules. However, this period has led to serious doubts on the risk that such UAV operations pose to manned aircraft operating at lower altitudes.

It must be underlined that the new regulation allows for BVLOS UAV operations outside of segregated airspace. Similar regulations were adopted in other European countries, but these include much stricter limitations. For instance, in France and Spain, such operations are only allowed for UAVs that weigh less than 2 kg [30]. In Spain, however, UAVs with a mass up to 25 kg can operate in BVLOS operation outside of segregated airspace when equipped with an approved detection and avoidance systems [31]. Similar regulations apply in the UK, where operators are required to implement an approved method of aerial separation and collision avoidance as a prerequisite for BVLOS operations [32]. Under Polish law, such systems are not considered as mandatory for BVLOS UAV operations outside of segregated airspace. And this is the main concern regarding the new regulation, even though it also creates the legal framework for the introduction of a UTM system [33].

While operating under national regulations (the details of which may vary between countries), UAV operators are required by law to be prepared for the new EU regulations that will apply across all Member States. The approach, terminology and the requirements used at an EU-level are quite different to national regulations as they are drafted for all EU Member States.

EU countries having their own rules are now obliged to make the transition from old regulations to new as smooth and seamless for drone operators as possible. One example is that the Civil Aviation Authorities will be obliged to transform the current UAV VLOS and BVLOS "licenses" into European equivalents, but it is currently unknown whether operators will be required to undergo supplementary training. The transformation of operator qualification certificates is not the only problem. Poland needs to decide whether new BVLOS regulations will stay in force for operations other than civil.

Due to the fact that EU law applies uniformly in all Member States, it is advisable to discuss the regulations that apply in non-EU countries, and to compile EU countries' own regulations to the extent that EU law allows the state's competence to regulate.

Below is a table that outlines how individual national regulations shape UAV flight and ownership in selected European Union countries.

As shown in the few examples included in Table 2, national regulations concerning UAV operations vary widely from country to country. These differences only hamper the growth of this sector and should be unified in accordance with the EUs Pan-European regulations for UAV operations explained earlier.

4. Results

The EU's new set of UAV rules for the European market which will facilitate the development of drone technologies and create plenty of job and business opportunities. Besides addressing vital issues such as citizen's privacy, security and environmental topics, a far-sighted vision sees this harmonization lead to the creation of a functioning UTM system through U-Space, enabling the large-scale performance of UAV operations in low-level airspace, beyond visual line of sight and in congested areas.

It is therefore critical that these new rules were created by EASA, as this organization is responsible for European air safety is capable of holding extensive stakeholder consultations.

Although it is too early to assess the new EU regulations in practice, since they will only enter into force later this year, this

Table 2: National regulations in selected EU countries

UAV Classification	UAV Requirements	Pilot Requirements	Permits
Austria – Austro Control [34]			
Austria classifies drones by joules of kinetic energy	1. No flying over crowds (e.g. events, sports events, concerts, etc.) without special permission 2. No flying in the immediate vicinity of airports without special permission	1. Drone pilots must be 16 years of age or older 2. Liability insurance is mandatory for all drone operators 3. Only Austro Control may grant an aviation permit to an operator	1. A permit is required if 79 joules are exceeded or if a drone is flown above 30 m 2. A permit is also required to fly a drone for commercial purposes or to take photos or record video regardless of operating mass
Belgium – Belgian Civil Aviation Authority (BCAA) [35]			
Belgium has three different classes of operations, with licensing requirements for each class	Class 1a and 1b Operations Definition: 1. Up to 300 ft (around 90 m) above ground outside of controlled airspace. 2. Operations can only occur in daylight conditions. 3. Drone must weigh less than 330 lbs. (150 kg). Class 2 Operations Definition: 1. Flying height limit of 150 ft (around 45 m) outside of controlled airspace and outside of cities or communities. 2. Operations can only occur in daylight. 3. The drone must weigh less than 11 lbs. (5 kg).	1. The pilot must maintain a visual line of sight with the drone at all times.	Licensing Requirements: 1. Class 1a and 1b Operations – require a remote pilot license issued by the BCAA. To obtain this license, UAV pilots need to pass both a theoretical examination and a practical skill test. 2. Class 2 Operations – UAV pilots need a certificate of competence issued by the BCAA. To obtain this certificate, UAV pilots need to follow a theoretical course and pass a practical skill test.
France – French Civil Aviation Authority (DGAC) [36]			
Commercial and Recreational	1. Drones may not be flown at night (unless with special authorization from the local prefect). 2. Drones may not be flown over people; over airports or airfields; over private property (unless with owner's authorization); over military installations, prisons, nuclear power plants, historical monuments, or national parks. Drones may also not be flown over ongoing fires, accident zones, or around emergency services. 3. Drones may not be flown above 150 m (492 ft), or higher than 50 m (164 ft) above any object or building that is 100 m (328 ft) or more in height.	1. Drone pilots must maintain a visual line of sight with their drones at all times. If a visual observer is tracking the drone, the pilot may fly out of his or her own range of sight.	1. Commercial – commercial drone pilots must pass a theoretical exam to receive a theoretical telepilot certificate. The pilot must have this printed and with them during all flights. Commercial drone pilots must also undergo basic practical training. Recreational – recreational drone pilots do not need a training certificate when UAV mass is less than 800 g. They must undergo training is UAV mass exceeds this. 2. All drones of 800 g or more must be registered by their owner on AlphaTango, a public portal for users of remotely piloted aircraft. The drone then receives a registration number that must be affixed on the drone.
Italy – Italian Civil Aviation Authority (ENAC) [37]			
Commercial and Recreational	1. Drones being flown for recreational purposes may not fly more than 70 m (230 ft) above the ground. For commercial purposes it is 150 m (492 ft). 2. Drones must be identified by a plate showing the identification of the system and of the operator. An identical plate shall be installed also on the remote ground pilot station. 3. UAVs must be equipped with an Electronic Identification Device. 4. Drones may not be flown within 5 km (3 miles) of any airport or airfield. 5. Drones may not be flown at night. 6. Drones are not allowed to fly over people or crowds, including sports events, concerts, and other large events.	1. Drone pilots must maintain a direct line of sight with their drone during operations.	1. Commercial drone pilots conducting low-risk operations must submit a statement of compliance with specific requirements to ENAC along with a 94 EUR processing fee. For higher-risk operations, commercial drone pilots must obtain training and an operating certificate as well as a health certificate.
Netherlands – Netherlands Directorate General of Civil Aviation (DGCA) [38]			
Commercial and Recreational	1. Flying a drone is only allowed in category G airspace. 2. Drones may not fly more than 120 m (394 ft) above the ground. 3. Drones must fly at a safe distance from people and buildings and may not be flown at night. 5. The maximum weight for private drones is 25 kg. 6. UAV pilots may use drones to make aerial films and photographs for their own personal use. This is regarded as recreational use.	1. Drone pilots must give priority to all other aircraft, such as airplanes, helicopters, gliders, etc. 2. Pilots must maintain a visual line of sight with their drone during operations.	1. Commercial drone operations in the Netherlands require the drone pilot to hold a pilot's license, and the company/organization overseeing the operation to hold am ROC permit to fly. 2. If UAV pilots operate a drone weighing no more than 4 kg, then they may apply for a light permit. 3. Drone insurance is required for commercial drone operations in the Netherlands.
Spain – Spanish Aviation Safety and Security Agency (AESA) [39]			
Recreational, Model Aircraft, Professional	1. Drones may be flown up to 120 m (394 ft) above the ground. 2. Drones may only be flown during the day. 3. Flights may be carried out at night for drones with a take-off weight of less than 2 kg (4.4 pounds) as long as a flight altitude of 50 m (164 ft) above the ground is not exceeded. 4. Drone pilots must maintain a distance of 150 m (492 ft) from buildings, and a distance of 50 m (164 ft) or more from people not involved in the flight. 5. Drone pilots must maintain a distance of at least 8 km (5 miles) to airports in uncontrolled airspace, or 15 km (9.3 miles) on approved BVLOS flights.	1. Drone pilots must maintain a visual line of sight with their drone at all times. 2. BVLOS allowed if a second visual observer monitors the drone and is in direct eye contact with the pilot	1. A permit is required for commercial drone flights. 2. Liability insurance is required for commercial drone pilots. 3. For flights in national parks, UAV pilots need permission from AESA. 4. The use of drones in no-fly zones must be approved by the Spanish Ministry of Defense.

study shows that certain EU countries (such as Poland) tried to stay ahead of the curve and adapt their national policies to developing and growing markets and technologies. Also, one of the greatest challenges lying ahead for all Member States is making the transition and replacing their current national rules – those that have been widely accepted by UAV operators in them – with an "unknown" set of regulations put forward by a supranational body. Adapting these to the specificity of each Member State's legal framework will certainly be a difficult task.

It must not be forgotten that despite having a unifying nature, the EU's regulations were created so quickly that they could not possibly cover all UAV-related issues, leaving many blank spaces to be filled in and questions to be answered by national legislators themselves.

5. Conclusion

As we enter the two-year transition period, EASA if aware of the difficulties that lie ahead. The agency even summed up at the recent High Level Conference on Drones, which was held during Amsterdam Drone Week this past December, that a common European vision in which UAVs can be used for emergency services, transportation or even parcel delivery will only become a reality if safety standards are met and solutions are found to pressing issues such as noise and privacy [40].

However, we can be quite certain that this unprecedented continental initiative to unify regulations will foster the growth of the UAV technology sector, allowing private businesses to drive the commercial market forward. May this example of European cooperation serve as a model for other nations around the world as we stand at the brink of a development that could revolutionize how drone technology is used across the region.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Research financed by National Science Center, Poland. Project No 2017/27/B/HS5/0008 Unmanned Aerial Vehicle. A New Era in Aviation Law.

References

[1] https://uavcoach.com/droneii-2019-reports/ [Accessed: 13-Jan-2020]
[2] https://www.droneii.com/the-drone-market-2019-2024-5-things-you-need-to-know [Accessed: 09-Jan-2020]
[3] https://www.tractica.com/newsroom/press-releases/commercial-drone-hardware-and-services-revenue-to-reach-12-6-billion-by-2025/ [Accessed: 09-Jan-2020]
[4] European Drones Outlook Study Unlocking the value for Europe. [Online]. Available: https://www.sesarju.eu/sites/default/files/documents/reports/European_Drones_Outlook_Study_2016.pdf, pp.3
[5] D. Chowdhury, M. Sarkar, M. Z. Haider "A Cyber-Vigilance System for Anti-Terrorist Drives Based on an Unmanned Aerial Vehicular Networking Signal Jammer for Specific Territorial Security" ASTES Journal Vol. 3, No. 3, 43-50 (2018)
[6] D. Chowdhury, M. Sarkar, M. Z. Haider, S. A. Fattah and C. Shahnaz, "Design and implementation of a cyber-vigilance system for anti-terrorist drives based on an unmanned aerial vehicular networking signal jammer for specific territorial security," 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka, 2017, pp. 444-448.
[7] S. A. Fattah et al., "An aerial landmine detection system with dynamic path and explosion mode identification features," 2016 IEEE Global Humanitarian Technology Conference (GHTC), Seattle, WA, 2016, pp. 745-752.
[8] European Drones Outlook Study Unlocking the value for Europe. [Online]. Available: https://www.sesarju.eu/sites/default/files/documents/reports/European_Drones_Outlook_Study_2016.pdf, pp.14
[9] https://www.sesarju.eu/U-space [Accessed: 02-Jan-2020]
[10] P. Kasprzyk and A. Konert, Unmanned aircraft. A new era for aviation. A new era for aviation law?, Coventry Law Review no 12 / 2018.
[11] See more in ICAO RPAS CONOPS (March 2017) and EUROCONTROL RPAS ATM CONOPS (Ed. 4.0, 2017)
[12] European Aviation Safety Agency, Regulatory framework to accommodate unmanned aircraft systems in the European aviation system, ISSUE 2 — 4.6.2018. [Online]. Available: https://www.easa.europa.eu/sites/default/files/dfu/ToR%20RMT.0230%20%E2%80%93%20Issue%202.pdf
[13] https://www.easa.europa.eu/newsroom-and-events/news/eu-wide-rules-drones-published [Accessed: 13-Jan-2020]
[14] https://www.easa.europa.eu/newsroom-and-events/press-releases/eu-wide-rules-drones-published [Accessed: 13-Jan-2020]
[15] COM/2014/0207 final
[16] Regulation (EC) No 216/2008 of the European Parliament and of the Council of 20 February 2008 on common rules in the field of civil aviation and establishing a European Aviation Safety Agency, and repealing Council Directive 91/670/EEC, Regulation (EC) No 1592/2002 and Directive 2004/36/EC (Official Journal of the European Union, L 79, 19.3.2008).
[17] P. Kasprzyk and A. Konert, Unmanned aircraft. A new era for aviation. A new era for aviation law?, Coventry Law Review no 12 / 2018.
[18] Commission Delegated Regulation (EU) 2019/945 of 12 March 2019 on unmanned aircraft systems and on third-country operators of unmanned aircraft systems https://eur-lex.europa.eu/eli/reg_del/2019/945/oj
[19] Commission Implementing Regulation (EU) 2019/947 of 24 May 2019 on the rules and procedures for the operation of unmanned aircraft https://eur-lex.europa.eu/eli/reg_irmpl/2019/947/oj
[20] https://www.easa.europa.eu/document-library/opinions/opinion-012018 [Accessed: 13-Jan-2020]
[21] https://www.airport-technology.com/features/new-eu-drone-rules/ [Accessed: 09-Jan-2020]
[22] https://dronerules.eu/assets/covers/Table-1.png [Accessed: 09-Jan-2020]
[23] https://www.commercialuavnews.com/europe/easa-publishes-common-drone-rules-for-all-of-europe [Accessed: 13-Jan-2020]
[24] https://dronerules.eu/en/professional/eu_regulations_updates [Accessed: 09-Jan-2020]
[25] Journal of Laws from 2002 No 130, Item 1112 with further amendments. See also M. Żylicz (ed.), Prawo lotnicze. Komentarz (Air Law Commentary). Warsaw 2016
[26] https://dronerules.eu/assets/images/DroneRules_v2.jpg [Accessed: 09-Jan-2020]
[27] Journal of Laws from 2013, Item 440
[28] Journal of Laws from 2016, Item 1317
[29] See more details in P. Kasprzyk & A. Konert, Drones are flying outside of segregated airspace in Poland – new rules for BVLOS UAVs operations, IEEE proceedings: https://ieeexplore.ieee.org/document/8798085
[30] LOI n° 2016-1428 du 24 octobre 2016 relative au renforcement de la sécurité de l'usage des drones civils (JORF n°0249 du 25 octobre 2016) (Law n° 2016-1428 of October 24th, 2016 on the reinforcement of the safety of the use of the civil drones)
[31] Real Decreto 1036/2017, de 15 de diciembre, por el que se regula la utilización civil de las aeronaves pilotadas por control remoto, y se modifican el Real Decreto 552/2014, de 27 de junio, por el que se desarrolla el Reglamento del aire y disposiciones operativas comunes para los servicios y procedimientos de navegación aérea y el Real Decreto 57/2002, de 18 de enero, por el que se aprueba el Reglamento de Circulación Aérea. (BOE Núm. 316 Viernes 29 de diciembre de 2017) (Royal Decree 1036/2017, of December 15, which regulates the civil use of drones and amending Decree 552/2014)
[32] CAP 722: Unmanned Aircraft System Operations in the UK Airspace – Guidance. [Online]. Available: https://publicapps.caa.co.uk/docs/33/CAP%20722%20Sixth%20Edition%20March%202015.pdf
[33] Such as "Drone Radar" DAMS (Drone Awareness and Monitoring System), see more on https://droneradar.eu/blog/droneradar-for-ansp/
[34] https://www.austrocontrol.at/jart/prj3/ac/main.jart?reserve-mode=active&rel=en. https://uavcoach.com/drone-laws-in-austria/ [Accessed: 14-Feb-2020]
[35] https://mobilit.belgium.be/fr/transport_aerien/drones, Drone operation type table: https://149355317.v2.pressablecdn.com/wp-content/uploads/2018/05/bcaa-drone-operations.png [Accessed: 14-Feb-2020]
[36] DGAC website: https://www.ecologique-solidaire.gouv.fr/direction-generale-laviation-civile-dgac, Vast legal report:

https://www.loc.gov/law/help/regulation-of-drones/france.php, AlphaTango website: https://fox-alphatango.aviation-civile.gouv.fr/en/, Drone law overview: https://dronerules.eu/assets/regulationspdfdownloads/NRP_france.pdf [Accessed: 14-Feb-2020]
[37] https://www.enac.gov.it [Accessed: 14-Feb-2020]
[38] Recreational drone use: https://www.government.nl/topics/drone/rules-pertaining-to-recreational-use-of-drones, Commercial drone use: https://www.government.nl/topics/drone/rules-pertaining-to-the-commercial-use-of-drones, DGCA website: https://www.ilent.nl/zoeken?trefwoord=drone [Accessed: 14-Feb-2020]
[39] https://www.boe.es/buscar/doc.php?id=BOE-A-2017-15721, https://www.seguridadaerea.gob.es/lang_en/cias_empresas/trabajos/rpas/default.aspx, https://dronerules.eu/assets/regulationspdfdownloads/NatinalRegulatoryProfile_Spain.pdf [Accessed: 14-Feb-2020]
[40] https://www.easa.europa.eu/newsroom-and-events/press-releases/regulators-and-industry-unite-need-address-societal-concerns [Accessed: 13-Jan-2020]

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 93-99 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Constrained Path Planning for Both Smooth Motion Profile and Stable Control

Anh Son Tran[1], Ha Quang Thinh Ngo[2,*]

[1]*Dept. of Manufacturing Engineering, HCMC University of Technology, Vietnam National University (VNU-HCM), Vietnam*

[2]*Dept. of Mechatronics Engineering, HCMC University of Technology, Vietnam National University (VNU-HCM), Vietnam*

ARTICLE INFO

Article history:
Received: 23 October, 2019
Accepted: 10 December, 2019
Online: 15 January, 2020

Keywords:
Polynomial Interpolation
Automation
Vibration-less

ABSTRACT

In our era, human and machine must share the same workspace in various plants such as warehouse logistics, material transportation, manufacturing factory or assisted human. The collaboration is essential to maintain the safe operation and the complete mission. In this paper, a design of anti-vibration scheme using Spline generator for grounded vehicle platform in human-machine interaction is introduced. This hardware structure is able to play a role as carrier in loading application or human assistant in real life. In most of cases, vehicle must convey cargo from start point to destination instead of human while its reference trajectory conflicts with human's movement. The proposed algorithm helps to carry out a decision of stopping strategy with the least vibration as possible to avoid accident. From the test results, the proposed design is feasible and capable to apply in numerous industries.

1. Introduction

In the first and second industrial revolution, human labor is the main factor to build the products in the workshops, factories, or warehouses. The productivity of this stage mostly depends on the human factor such as clock operation, time off, overtime, short-term payment. Until 1950s, the Automated Guided Vehicle (AGV) was firstly introduced as an advanced mobile robot which uses wires, UV marks and light to navigate. Nowadays, AGVs are even more advanced by using magnets, laser range finder or digital cameras to pilot. Moreover, the most complex guidance technique is currently an integrated on-board with sensors and software, using wireless or radio frequency connection for routing. This autonomous system is commonly applied in large scale factories and storehouses to transport materials from one place to another or assist human by following them with packages. AGVs come with powerful computers that are connected to the cameras, lasers and sensors. This allows vehicle to handle multi-tasking or complex missions all the time. It is completely safe to walk or stand around AGVs while they are operating. On the other hand, unlike human, AGVs can work overtime and will result in the growth of productivity and minimize the expenses.

However, in most of the autonomous system, while they are working, there are always exist residual vibration, which leads to the lack of accuracy. For instance, the crane with heavy load in the end-effector is moving from one place to another. And when it's stopped, the vibration occurs proportionally to the mass of the load, which not only lead to the inaccuracy but also damage the joints of the crane. In other example, the AGV is following human with a bulky material above and still maintain the safe distance to human. It is assumed that sudden collision would force AGV to stop immediately. If it is not stopped properly, the materials above would collapse due to the vibration. In [1, 2], the researchers composed of four major functional blocks including human detection and human feature, dynamic social zone, approaching pose of the robot to a human or a human group. They are incorporated into a motion planning system, comprising a local path planner and dynamic window technique. The output of research is only focus on driving a mobile robot to approach both stationary and moving humans in a socially acceptable manner. They did not concern on dynamic constraints of system model to guarantee the safe operation. In the applications where the oscillating dynamics is not easily detectable with the standard sensors that hung on the plant, accordingly, it cannot be compensated by a feedback control action. The use of a feed-forward control is often the only option for reducing the vibrations. this technique [3-5] is very effective in vibration suppression, but it may be not useful in trajectory following application such as 3D printing, cutting ... The alternative method to dominate oscillation

* Ha Quang Thinh Ngo, Email: nhqthinh@hcmut.edu.vn

is the use of finite impulse response filter for planning minimum-time trajectories [6, 7] was presented. This investigation connects between analytic functions and dynamic filters permit a generalization of various trajectories. It is usually obtained by second- or third-order polynomial functions to a generic order with only a modest increase of the complexity. The exponential jerk along with filters was investigated as a new type of trajectory [8, 9]. As a matter of fact, constant jerk trajectories did not guarantee a complete vibration suppression when the damping of the resonant modes is not negligible. The decay rate and duration of the jerk impulses that allow residual vibration cancellation are derived in an analytical way as a function of the dynamic characteristics of the plant. Some comparisons with the well-known input shaping techniques combining with filters performs the advantages of this method. In [10-12], the authors demonstrated the theories of motion profiles from basic to advance, the mechanism of motion generators and their applications in industry. They acquainted with polynomial profile, exponential profile and multi-segmented profile to figure out the generation of robot trajectory. One of advanced path generators is a uniform B-splines of degree p. The investigators in [13-15] developed the spline trajectories with a chained composed by p average filters to optimize. It could find out the solutions for two different troubles such that exactly interpolating a set of given points by means of a complex trajectory and the need of suppressing mechanical vibrations. However, these researches are only simulated in virtual environment from theories. It needs more time to overcome practical problems, for example, time consuming to generate motion profile, burden computation in computer or high cost comparing to other methods.

In this paper, the exponential B-spline trajectories are taken into consideration and able to generate exactly if the profiling parameters are chosen properly. Besides, for safety in the working environment, there are many options such as stereo cameras, laser sensors, proximity sensors that could be placed additionally. A Kinect camera is chosen because of its advantages, i.e. cheap price, easy to calibrate, library supported and most importantly human detectable. A Kinect camera is attached on the vehicle to calculate the distances when humans are detected while moving along its path, based on the depth sensor integrated in the camera. The experimental results performed on an automatic guided vehicle, show the effectiveness of the proposed method.

2. B-spline and Applications

Splines are piecewise polynomial functions widely used to interpolate sets of data points or to approximate functions, curves and surfaces. The data may be either one-dimensional or multi-dimensional. A spline of order p is a piecewise polynomial function of degree $p-1$ in a variable of t. The values of u where the pieces of polynomial meet are known as knots, denoted $t_0, \dots, t_{knot}$ and sorted into non-decreasing order. When the knots are distinct, the first $p-1$ derivatives of the polynomial pieces are continuous across each knot. When r knots are coincident, only $p-r$ derivatives of the spline are continuous across that knot.

2.1. B-spline Functions

A particularly efficient technique for the computation of splines is based on so called B-splines, or Basic-splines. The reason for such a name is that a generic spline can be obtained as a linear combination of a proper number of B-spline basis functions, $B_j^p(t)$, i.e.

$$s(t) = \sum_{j=0}^{m} p_j B_j^p(t) \tag{1}$$

where

$s(t)$ is the curve trajectory.

$p_j, j = 0, \dots, m$ are the control points.

$B_j^p(t)$ is the jth B-spline basis function of degree p.

$t = [t_0, \dots, t_{knot}]$ is the knot vector.

B-spline basis function can be generated by using Cox-de Boor formula or convolution method. B-spline is a general term.

where

$B^{(0,0,0,0)}(t) = B^3(t)$ is a cubic polynomial spline.

$B^{(0,0,0,\widehat{\alpha})}(t)$ is an exponential spline of degree 3.

$B^{(0,0,\widehat{-\alpha},\widehat{\alpha})}(t)$ is an exponential tension spline of degree 3.

$B^{(\widehat{\alpha_0},\widehat{\alpha_1},\widehat{\alpha_2},\widehat{\alpha_3})}(t)$ is an exponential general spline form produces B-splines that are a linear combination of function $e^{\widehat{\alpha_0}t}, e^{\widehat{\alpha_1}t}, e^{\widehat{\alpha_2}t}, e^{\widehat{\alpha_3}t}$.

Cubic polynomial spline can be generated in both methods. Exponential spline can only be generated by using convolution method.

Cubic polynomial spline basis function using Cox-de Boor. In this method, B-splines are piecewise polynomial, characterized by a knot vector $t = [t_0, \dots, t_{knot}]$. The relation between knots, n_{knot} and number of control points, m is $n_{knot} = m + p + 1$.

$$B_j^0(t) = \begin{cases} 1 & t_j \le t < t_{j+1} \\ 0 & otherwise \end{cases} \tag{2}$$

$$B_j^p(t) = \frac{t - t_j}{t_{j+p} - t_j} B_j^{p-1}(t) + \frac{t_{j+p+1} - t}{t_{j+p+1} - t_{j+1}} B_{j+1}^{p-1}(t) \tag{3}$$

Exponential spline basis function using convolution. In this method, B-splines characterized by an equally-spaced distribution of the knots $t_{j+1} - t_j = T, j = 0, \dots, m-2$.

$$B^{\alpha_k}(t) = \begin{cases} \dfrac{\alpha_k T}{e^{\alpha_k T} - 1} e^{\alpha_k t} & 0 \le t < T \\ 0 & otherwise \end{cases} \tag{4}$$

$$B^{(\alpha_0,\alpha_1,\alpha_2,\alpha_3)}(t) = \frac{1}{T} B^{\alpha_3}(t) * \dots * \frac{1}{T} B^{\alpha_1}(t) * B^{\alpha_0}(t) \tag{5}$$

2.2. B-spline Applications

The linear combination of B-spline basis functions with a set of control points can solve the interpolation problem, therefore, in mathematical subfield of numerical analysis, it is widely used for curve fitting and numerical differentiation of experimental data. In computer-aided design and computer graphic, B-spline basis functions are used for generating and representing curves or surfaces.

3. Motion Planner Using B-Spline

Smoothness, flexibility and possibility of exactly interpolating some via-points are the main characteristics of B-spline functions. Therefore, we believe that it is possible to generate B-spline trajectories which exactly cross the desired via-points and at the same time are able to suppress vibrations.

3.1. B-spline Curve Generator

Using Cox-de Boor formula, it can only solve the cubic polynomial spline case. Though, it's a practical method to generate the B-spline trajectory for autonomous system while maintaining the properties of system. The proposed algorithm of flexible and controllable motion generator is illustrated in Fig. 1. In this case, the B-spline curve has degree 3 together with its control polygon defined by.

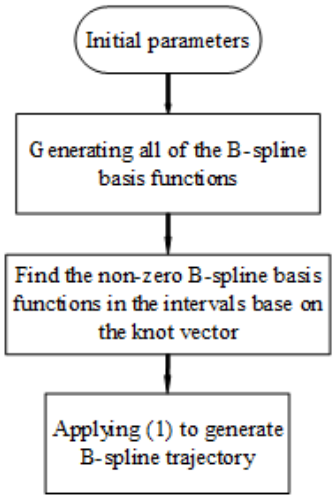


Figure 1. Proposed algorithm of B-spline trajectory generator.

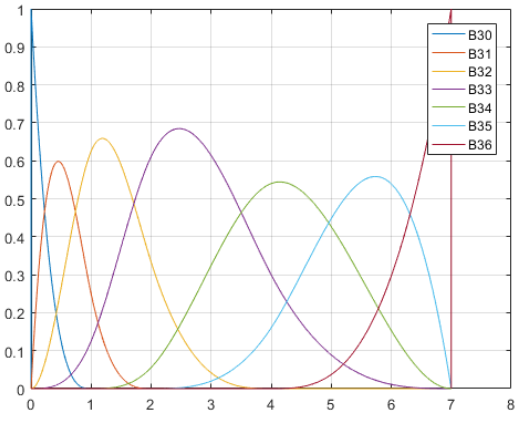


Figure 2. B-spline basis functions using Cox-de Boor formula

$$P = \begin{bmatrix} p_0, & p_1, & \cdots & p_{m-1}, & p_m \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & -3 & 4 & 5 & -5 & -6 \end{bmatrix} \quad (6)$$

The knot vector is,

$$u = \begin{bmatrix} 0, & 0, & 0, & 0, & 1, & 2, & 4, & 7, & 7, & 7, & 7 \end{bmatrix} \quad (7)$$

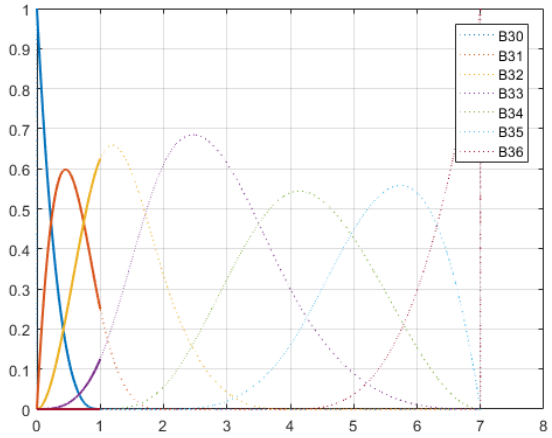


Figure 3. Non-zero B-spline basis functions in the interval [0,1].

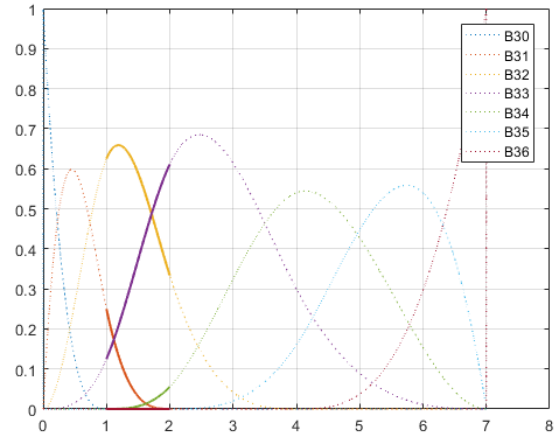


Figure 4. Non-zero B-spline basis functions in the interval [1,2].

To generate the B-spline with above parameters, a method of path planning using various basis functions has been investigated. According to Fig. 1, with the initial parameters (the control point vector and the knot vector) and a combination with (3), it is able to generate the B-spline basis functions, note that the variable t_j in (3) now becomes u_j in (7) with $j = 0, \ldots, 10$ with respect to the example. Since it has a total of seven control points, it requires seven B-spline basis functions of degree 3 which are $B_0^3(t), \ldots, B_6^3(t)$ to be able to calculate the trajectory using (1). After calculating all of the B-spline basis functions, note that in every knot span $[u_j, u_{j+1})$, there are at most $p + 1$ basis functions B_j^p are not null, namely $B_{j-p}^p, \ldots, B_j^p$, for example there are 4 basis functions $B_0^3, B_1^3, B_2^3, B_3^3$ are not null in the interval $[u_3, u_4)$ as shown in Fig. 5. Using (1), for each of the B-spline basis functions of degree 3 as mentioned above, multiply it with the corresponding control points, as a result we have $s(t)$, which is a vector with the first row stands for the time in the x-axis and second row stands for the amplitude of the trajectory in the y-axis.

Fig. 2 introduces all of the B-spline basis functions in the interval [0,7). Note that, at every sampling time, the sum of all of

the B-spline basis functions at that time is always equal to 1, this is also one of the characteristics of the B-spline basis functions.

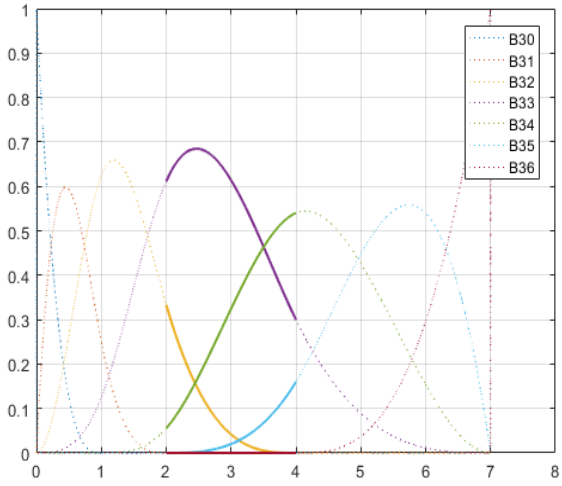


Figure 5. Non-zero B-spline basis functions in the interval [2,4].

In Fig. 3 and Fig. 4, there are 4 non-zero basis functions which are available in the interval $[0,1)$ and $[1,2)$, they are $B_0^3, B_1^3, B_2^3, B_3^3$ and $B_1^3, B_2^3, B_3^3, B_4^3$ respectively.

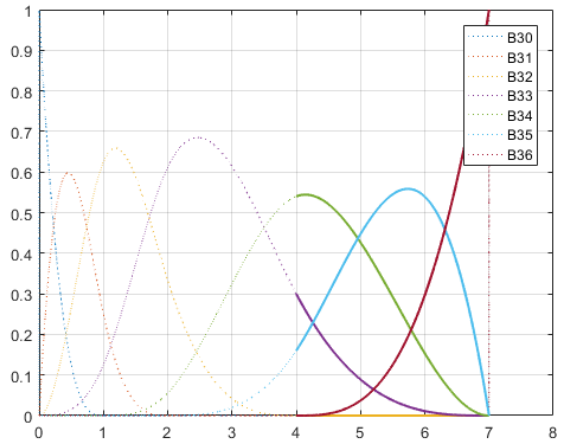


Figure 6. Non-zero B-spline basis functions in the interval [4,7].

3.2. Smooth Profiling Generation

In order to prove the effectiveness and correctness of B-spline application for vibration-less motion, a practical example of material transportation in factory has been chosen. An automated guided vehicle (AGV) with cargo must travel in long distance during working hours. It is recognized that the stable movement is difficult to gain since the vibration occurs. To conquer these troubles, we apply B-spline trajectory to reduce the oscillation. Fig. 10 illustrates the system modeling of both AGV and cargo as m_1 and m_2. They are linked together via a string (stiffness k) and damper (viscous coefficient b).

$\vec{F}_{qt}$: inertial force acting on m_2

$\vec{v}$: velocity of m_1

$\vec{x}_1, \vec{x}_2$: displacement of m_1 and m_2

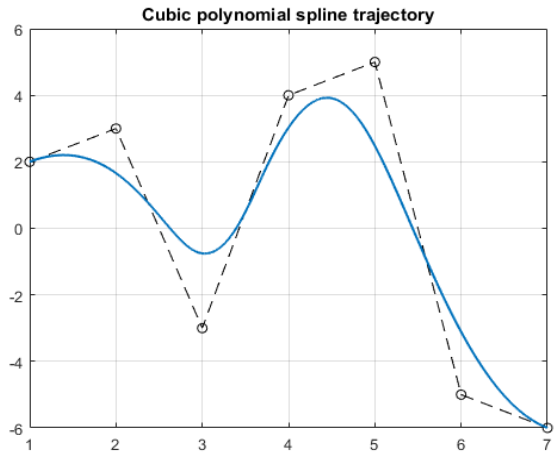


Figure 7. Cubic polynomial spline trajectory.

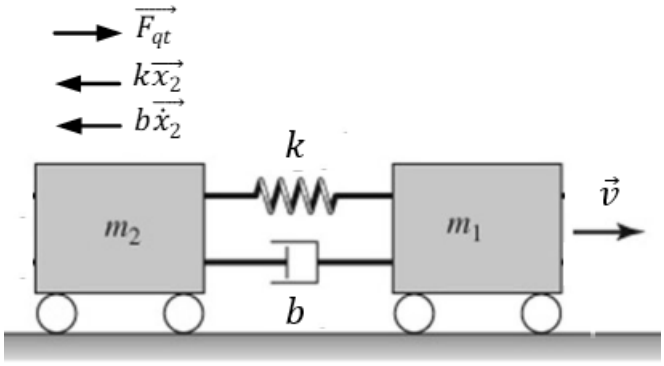


Figure 8. Modeling of both vehicle and cargo.

Using Newton law, we have equation of system dynamics in time domain

$$m_2\vec{a}_2 = \vec{F}_{qt} + k\vec{x}_2 + b\dot{\vec{x}}_2 \quad (8)$$

In its direction,

$$m_2\ddot{x}_2 = -kx_2 - b\dot{x}_2 + m_2\ddot{x}_1 \quad (9)$$

By applying Laplace transformation, the below description is in frequency domain

$$m_2 s^2 X_1 = m_2 s^2 X_2 + kX_2 + bsX_2 \quad (10)$$

$$\Leftrightarrow X_2 = \frac{s^2 X_1}{s^2 + 2\varepsilon\omega_n s + \omega_n^2} \quad (11)$$

$$\Leftrightarrow \frac{X_2}{X_1} = \frac{s^2}{s^2 + 2\varepsilon\omega_n s + \omega_n^2} \quad (12)$$

where,

$\omega_n = \sqrt{\frac{k}{m_2}}$: the natural frequency of the of the lightly damped system.

$\varepsilon = \frac{b}{2m_2\omega_n}$: the damping ratio.

From equation (11), it has been known that the system vibrates more at pole points while it maintains a stable state at zero points are the points. On the other hand, the system is stable if the poles lay on the left side of the complex plane or being replaced by the zeros using the pole placement technique. In (12), we can find the zeros and poles of the system.

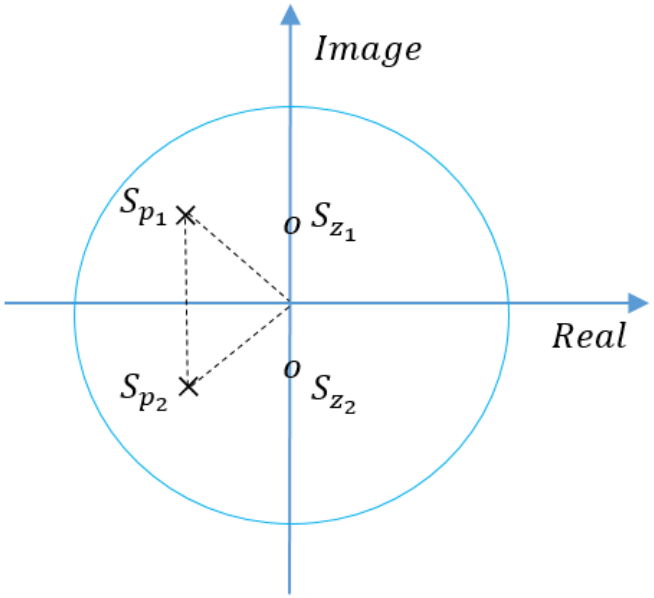


Figure 9. Poles and Zeros in the complex plane.

Zeros,

$$s^2 X_1 = 0 \quad (13)$$

$$\leftrightarrow s_{z_{1,2}} = \pm j\frac{2\pi k}{t_3 - t_2} \quad (14)$$

Poles,

$$s^2 + 2\varepsilon\omega_n s + \omega_n^2 = 0 \quad (15)$$

$$\leftrightarrow s_{p_{1,2}} = -\varepsilon\omega_n \pm j\omega_n\sqrt{1-\varepsilon^2} \quad (16)$$

The zero points together with pole points are illustrated in the complex plane as Fig. 9 below.

4. Planning the Proposed Scheme

In the present work, we discover the case of automated vehicle which lifting cargo to transport material in factory. The test scenario was set up as Fig. 10 which the proposed scheme is experimented on a model of vehicle. It is attached the flexible beam together with mass and an accelerometer above. In front of vehicle, one digital camera is hung on anterior cover. The distance to conflict with objects could be estimated since it returns all coordinates in workspace. The test surrounding is internal area of textile factory where there are a lot of workers. As a result, the potential accident between human and machine could happen at any time. The proposed scheme is able to overcome these troubles perfectly.

In detail, the automated guided vehicle is driven by using a TI's board to output PWM signal. A powerful DC driver that is capable to protect the over current or isolate the reverse current, is motivated to control DC motor. The accelerometer is a Sensors Booster-Pack, manufactured by Texas Instrument and include several sensors into one pack. We utilize this sensing module to recognize the magnitude of oscillation in respect to time. This sensor would direct data back to host server through UART protocol. Whenever the mobile vehicle moves, with mass, the oscillation occurs to act on aluminum beam. The swinging of beam is recorded by sensing module. Later, oscillating data is transmitted to host personal computer. Moreover, a Kinect camera locating ahead of robot, has fully functional image processing techniques. We program an interface software with camera in C/C++ language.

Theoretically speaking, it is desired that there exists an effective and feasible method to reduce the residual vibration in such situation. To verify our approach, the polynomial cubic spine, exponential spline of degree 3, exponential tension spline of degree 3 and exponential general spline scheme are recommended to achieve free-vibration motion. The geometric trajectory is collected via linear lines with circular lines in a sequence of equally segment due to control points p_i. Once, the flexible curves are generated to evaluate the exactly tracking ability of system and stable response. However, owing to residual vibration, the actual performance differs with command curvatures of the bends. For such fluctuation, it is traditionally considered that robot's direction movement along with y-axis would produce the residual vibration concerned on x-axis response. The higher velocity the autonomous robot reaches to, the larger amplitude of oscillation takes place.

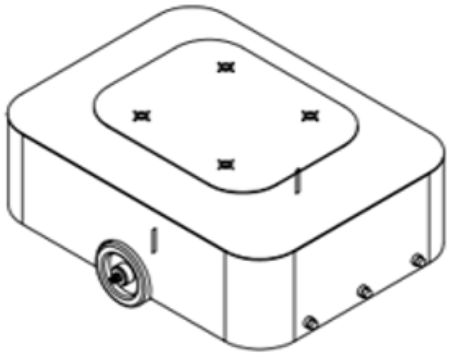


Figure 10. Simulation model of autonomous vehicle.

Based on proposed approach, an experimental prototype of potential collision has been built as Fig. 11. The differential drive mechanism depending on dissimilar speeds, orientates itself to turn left or right.

5. Discussion

The path planner based on cubic spline curve is denoted as Fig. 12. A series of B-spline basis functions are synthesized to perform the flexible and bended trajectory. Nevertheless, at each sampling time, there is only one basic function that is the most weighted scheme. Also, the determination of basic functions depends on control way-points. As we expect, the routing passes through all the given way-points.

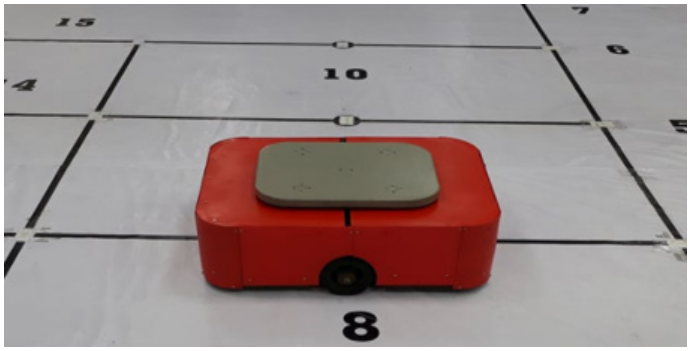


Figure 11. Practical model of autonomous vehicle.

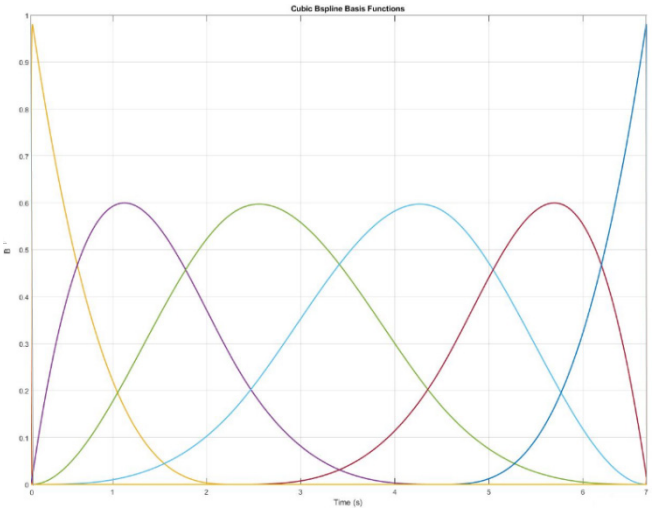


Figure 12. Cubic spline trajectory generation.

Note that there is a unique synthetization of control points which meets the way-point constraints. It could offer more degrees of freedom by swelling the number of control points. To utilize this generation more effective and real-time cost, a powerful CPU needs to be explored. The B-spline based motion profile of mobile vehicle is achieved as Fig. 13. The motion generator outputs driving command owing to interpolation functions. It provides an exact characterization of such category of routing line. While moving with feasible profile, the autonomous robot could ensure to reach desired parameters such position, speed, acceleration and jerk. To measure the impact of proposed approach, the reaction of mass m_2 is recorded in Fig. 14. The variation of m_2 tends to zero when time is infinite. From these results, it can be considerably seen that the path generation using B-spline strategy in respect to analysis of dynamic characteristics has been developed successfully. The wide range of industrial applications are recommended this study in future.

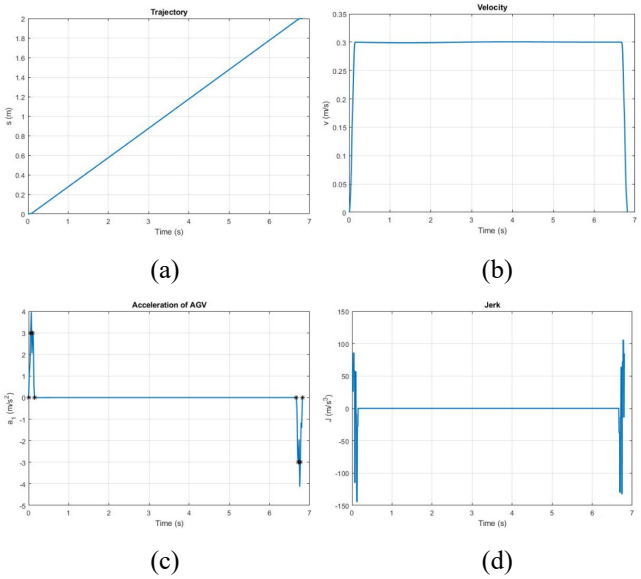


Figure 13. Motion parameters of autonomous vehicle, (a) traveling distance, (b) velocity, (c) acceleration and (d) jerk.

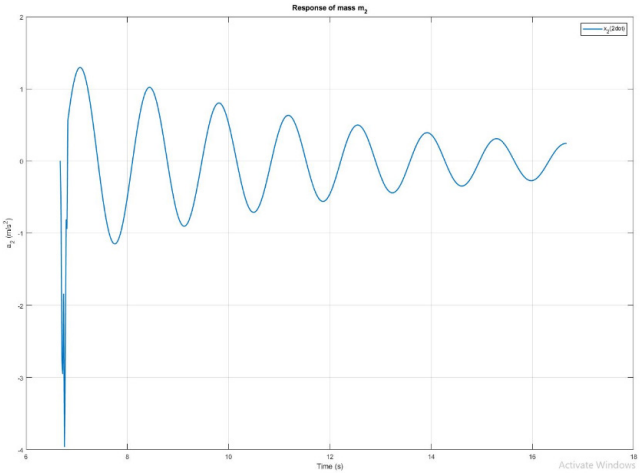


Figure 14. Actual response of carrier m_2.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] X. T. Truong and T. D. Ngo, "To Approach Human?: a unified framework for approaching pose prediction and socially aware robot navigation," IEEE Transactions on Cognitive and Developmental Systems, vol. 10, no. 3, pp. 1-17, 2018, https://doi.org/10.1109/TCDS.2017.2751963.

[2] F. Yang and C. Peters, "Social-aware navigation in crowds with static and dynamic groups," 11th International Conference on Virtual Worlds and Games for Serious Applications, pp. 1-4, 2019, https://doi.org/10.1109/VS-Games.2019.8864512.

[3] W.E. Singhose and N.C. Singer (1996), "Effects of input shaping on two-dimensional trajectory following," IEEE Trans. on Robotics and Automation, 12(6), https://doi.org/10.1109/70.544771.

[4] Q. T. Truong, H. Q. T. Ngo, T. P. Nguyen, H. Nguyen and W. H. Kim," A Novel Infrastructure Design of Industrial Autonomous System," International Journal of Fuzzy Logic and Intelligent Systems, 19(2), pp. 103-111, 2019, http://doi.org/10.5391/IJFIS.2019.19.2.103.

[5] H. Q. T. Ngo, Q. C. Nguyen, W. H. Kim, " Implementation of Input Shaping Control to Reduce Residual Vibration in Industrial Network Motion System," 2015 15th International Conference on Control, Automation and Systems, pp. 1693-1698, 2015, http://doi.org/ 10.1109/ICCAS.2015.7364629

[6] L. Biagiotti and C. Melchiorri (2012), "FIR filters for online trajectory planning with time- and frequency domain specifications," Control Engineering Practice, vol. 20, no. 22, pp. 1385-1399, 2012, https://doi.org/10.1016/j.conengprac.2012.08.005.

[7] H. Q. T. Ngo, T. P. Nguyen, T. S. Le, V. N. S. Huynh and H. A. M. Tran, "Experimental Design of PC-based Servo System," International Conference on System Science and Engineering, pp. 733-738, 2017, https://doi.org/10.1109/ICSSE.2017.8030973 .

[8] L. Biagiotti, C. Melchiorri and L. Moriello (2015), "Optimal trajectories for vibration reduction based on exponential filters," IEEE Trans. on Control Systems Technology, vol. 24, no. 2, pp. 609-622, 2016, https://doi.org/10.1109/TCST.2015.2460693 .

[9] H. Q. T. Ngo and M. H. Phan, "Design of an Open Platform for Multi-Disciplinary Approach in Project-based Learning of an EPICS Class," Electronics, 8(2), 200, https://doi.org/10.3390/electronics8020200.

[10] L. Biagiotti and C. Melchiorri, Trajectory Planning for Automatic Machines and Robots, Springer-Verlag Berlin Heidelberg, 2008.

[11] H. A. M. Tran, H. Q. T. Ngo, T. P. Nguyen and H. Nguyen, "Implementation of vision-based autonomous mobile platform to control by A* algorithm," International Conference on Recent Advances in Signal Processing, Telecommunications and Computing, pp. 39-44, 2018, http://doi.org/10.1109/SIGTELCOM.2018.8325802.

[12] N. M. Phu, V. Tuyen, T. T. Ngo, "Augmented heat transfer and friction investigations in solar air heater artificially roughened with metal shavings," Journal of Mechanical Science and Technology 33, 3521-3529 (2019), http://doi.org/10.1007/s12206-019-0646-x

[13] L. Biagiotti and C. Melchiorri (2011), "Input shaping via b-spline filters for 3-d trajectory planning," In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, http://doi.org/10.1109/IROS.2011.6094836.

[14] N. M. Phu, N. T. M. Trinh, "Modeling and experimental validation for off-design performance of the helical heat exchanger with LMTD correction taken into account," Journal of Mechanical Science and Technology 30, 3357-3364 (2016), http://doi.org/ 10.1007/s12206-016-0645-0

[15] H. Q. T. Ngo, Q. C. Nguyen and W. H. Kim, "Implementation of input shaping control to reduce residual vibration in industrial network motion system," International Conference on Control, Automation and Systems, pp. 1693-1698, 2015, http://doi.org/10.1109/ICCAS.2015.7364629.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 99-107 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Word Spotting Method for Arabic Manuscripts Based on Speeded Up Robust Features Technique

Noureddine El Makhfi*

Faculty of Science and Technology, Department of Mathematics and Computer Science, PMIC Lab, 32003, Al-Hoceima, Morocco

ARTICLE INFO

Article history:
Received: 15 August, 2019
Accepted: 07 November, 2019
Online: 22 November, 2019

Keywords:
Arabic calligraphy
Arabic handwritten documents
Cultural heritage
Digital libraries
Features extraction
Text-Image format
Segmentation
Word matching
Word spotting

ABSTRACT

The diversity of manuscripts according to their contents, forms, organizations and presentations provides a data-rich structures. The aim is to disseminate this cultural heritage in the images format to the general public via digital libraries. However, handwriting is an obstacle to text recognition algorithms in images, especially cursive writing of Arabic calligraphy. Most current search engines used by digital libraries are based on metadata and structured data manually transcribed in Ascii format. In this article, we propose an original method of pattern recognition for searching the content of Arabic handwritten documents based on the Word Spotting technique. Our method is both effective and simple, it consists in extracting a set of features from the words we segment in the target images and comparing them with the features of the words in the requested images. The principle of the method is to characterize each word with the Speeded Up Robust Features algorithm whose goal is to find all occurrences of query words in the target image even in the case of low-resolution images. We tested our method on hundreds of pages of Arabic manuscripts from the National Library of Rabat and the Digital Library of Leipzig University. The results obtained are encouraging compared to other methods based on the same Word Spotting technique.

1. Introduction

This article is an extended version of an oral communication presented in 2019 at the 5th International Conference on Optimization and Applications (ICOA2019) [1].

Virtual libraries, developed by research institutions, are a means of preserving documentary heritage in the form of images. The old manuscripts preserved in these libraries are consulted by a minority of people because of the image content that is inaccessible by search engines. Some techniques used for indexing manuscript images include metadata, annotation and transcription.
Metadata provides a means of describing the content of manuscripts in image form. On the other hand, the transcription or annotation of manuscripts offers an alternative to partially deepen the search in the content of the images. However, these techniques are very limited in requiring a considerable effort to reproduce the content of images in text format (ASCII) and do not offer a thorough and complete search. We find that copiers use the same writing style in old manuscripts. Each word or element (letter, pseudo word, calligraphy, etc.) is repeated several times in the corpus of Arabic manuscripts. In the absence of photocopiers at that time, most Arabic manuscripts are reproduced several times by the same copiers with the same styles of Arabic calligraphy (Naskh, Kufi, Diwani, Thuluth, Roqaa, etc.)

The idea is to reduce the number of elements to be transcribed manually when entering the content of images in Arabic manuscripts that use the same writing style. Our objective is to compare elements (content) of Arabic manuscripts based on a robust method. Comparing the content of manuscript images requires a multi-step process, each word or element being characterized by a unique signature. This signature must be invariable and must adapt to changes in scale, rotation, geometric variations and illumination. The extraction of the features of words or elements in the form of a signature, thus facilitates the automation of annotations or the complete transcription of Arabic manuscripts. Several research studies have been carried out in the field of access to the content of manuscripts in image form. However, most of this work is based on binary images, which results in a huge loss of information in the images. To overcome this problem, we used grayscale images to extract useful features in Arabic manuscript images. To this end, we have based ourselves on multi-scale theory using grayscale images to ensure maximum reliability when extracting all the features of the segmented words.

*Noureddine EL MAKHFI, Email: n.elmakhfi@gmail.com

In this paper, we present an original method for identifying the content of Arabic manuscript images. This method is based on the Word Spotting technique which allows the search for similar words in a target image to that of a query image. We used the Speeded Up Robust Features algorithm (SURF) [2]. This algorithm allows to extract the features of the images based on local invariant descriptors / detectors (at scale changes, rotation, luminosity variations, etc.). The SURF algorithm is known in an area of computer vision with its robustness and speed to detect and describe points of interest in 2D images. Arabic manuscripts presented in image form contain information that is repeated several times. The Word spotting technique we use allows us to search for similar words in the same manuscript in order to facilitate its identification, indexing or transcription.

2. Related Work

There are several recent projects, allowing the cataloguing and digitization of ancient manuscripts based on the metadata technique. We can mention the project called "Cultural Heritage" of the Digital Library of the Leipzig University [3]. This is a project aimed at the online dissemination of Arabic, Turkish and Persian Islamic manuscripts. The presentation of these manuscripts is in the form of high definition images, in different formats. Indexing and searching in the databases of this digital library are in Trilingual services (German, English, Arabic).
Most of the current research work is focused on researching the content of manuscript images. Among the most commonly used methods, the Word Spotting technique which allows the identification of words in images. There are recent studies that address the problem of indexing manuscripts and the recognition of cursive writing, namely:

In several studies [4-9], the authors treat the Word Spotting technique as regions of interest composed of words in which each character follows its model. The indexing and recognition of words are done by a probability calculation of regions of interest formed by characters. Each candidate region of interest is classified using one of the following methods: Dynamic Time Warping (DTW) or Hidden Markov Models (HMM) [10]. Despite the possible changes in words caused by character models, the results are interesting. Another work [11] which concerns the search for words based on annotations in medieval manuscripts. The techniques used are the basis of the Neural Networks (NN) and Hidden Markov Models (HMM) methods. The major problem with this method is the manual aspect of the production of annotations or transcription.

Other works that deal with word search based on the measure of similarity. The words, images are characterized according to their shapes using generic descriptors [12]. Among the shape descriptors that is used for word detection we can mention Histogram of Oriented Gradients (HOG) [13]. This description provides a means to describe and detect characteristics of handwritten words in [14].

In recent works [15-19], word detection is based on the SIFT algorithm [20,21]. This SIFT descriptor uses points of interest to describe the shapes of the words to be detected. The main problem with this technique used for word detection is the execution time and the false matching of the interest points of the images. Despite the excellence of the method for characterizing image words, like what is done in computer vision for the detection of 3D objects. The limitations of this method for word detection in 2D images can be improved by inserting other steps such as pre-processing.

In the general case, methods for detecting objects in images are represented by several research projects, particularly those in the field of image processing and computer vision. These methods are classified according to three approaches: contour, model-based and local extrema detection.

The first approach provides a means to detect storytellers based on the location of the shapes of the objects in the image. This approach exploits the geometric properties of objects in the image as corners to detect storyteller points. The second model-based approach provides a mapping of a theoretical model to points of interest in the intensity function. The latter approach provides a means to characterize and detect local extrema. Each extremum corresponds to points of strong intensity variations in the image.

2.1. Contour approach

The H. Moravec [22] detector is one of the older ones used for the interest points. This type of detector is highly dependent on containers and its major problem is its sensitivity to noise.

An advanced version of the H. Moravec detector by C. Harris and M. Stephen [23] is named the Harris detector. It's a corner detector. It is invariable to translation, to rotation. However, it does not provide good results in the case of change of illumination. Further improvements to the Harris detector were made by Mohr and Schmid [24]. P. Montesinos [25] has developed this detector to process even colour images. Other improvements were made by Y. Dufournaud [26] to make this detector invariable to changes in resolution. A highly advanced version of the Harris detector was developed by Mikolajczyk [27] to detect points of interest in scale space. This method consists of using Gaussian smoothing to get the points that have a local maximum in the scale space. A further improved version of the Harris detector that is invariable to affine transformations was developed by Harris-affine [28].

2.2. Theoretical signal model approach

The theoretical signal model approach was developed by Schmid [22]. It consists in bringing together shapes of objects in images with theoretical signal models to obtain more sub-pixellic precision. This approach allows to locate points in the shapes of complex objects and not only for corner detection. Rounded shapes such as circles and curves present an obstacle for corner detectors. Rohr [29, 30] developed a blur model based on the convolution of a Gaussian with a binary model to detect the junctions of lines. This model is based on several parameters for corner angles, namely the orientation of the symmetry axis and the opening. This model is also based on dot position, grayscale and blur. The adjustment of these parameters allows to obtain a theoretical signal approximating the observed signal in such a way as to obtain a result closer to the real. Corner detection is based on the least squares method in order to minimize parameters by offering a more precise model.

R. Deriche and T. Blaszka [31] optimized the processing time with the improvement of the Rohr method. The principle of their methods is based on the exponential function instead of using the Gaussian function. P. Brand and R. Mohr [32] improved

localization quality based on a theoretical model close to the signal by fine transformation. This reconciliation is performed with alignment tests, epipolar geometry calculation and 3D reconstruction. The authors announce that test results can achieve an accuracy of 0.1 pixels.

G. Giraudon and R. Deriche [31, 33, 34] proposed a theoretical corner model to increase detection accuracy. This model is used to track the behaviour of each detector. The authors demonstrated the relationship between responses in scale space and the position of the actual characteristic.

2.3. Approach based on local extrema

The approach based on local extrema allows to locate points of interest in the corresponding scale space. Several detectors have been processed in the literature and are widely used to locate local extrema. We quote some methods for detecting points of interest to characterize objects in the image:

T. Lindeberg [35] among the first who worked on scale space. The automatic scale detection method is used to characterize the points detected at the corresponding scale. The detected points are located at the maximum magnitude of the Laplacian. They correspond to the peak points of the scale characteristic of the associated local surface. This scale is determined in the main direction of dominant curvature.

The Scale Invariant Feature Transform (SIFT) method is presented by Lowe in [20, 21]. This method is based on the detection of points of interest corresponding to the local extrema. The principle of the method is based on the differences of the Gaussian DOG to detect invariable robust points in corresponding scales.

A study realised by Mikolajczyk and Schmid [36], proposed a method for the detection of invariable points of interest called Harris-Laplace. It is based on the former Harris detector, thus providing a localization of the local extrema. This method allows to keep only the local extrema at the neighbourhoods of the characteristic ladders in the form of invariable robust points of interest.

Another study realised by H. Bay et al [3], proposed a point of interest detection method based on the Speed Up Robust Features (SURF) algorithm. The method is also based on the location of local extrema as the other methods mentioned above: SIFT [21] by Lowe and Harris-Laplace [36] by Mikolajczyk. The authors of the SURF method have shown that their algorithm is more powerful, more robust. SURF's response time is largely optimized compared to the two SIFT and Harris-Laplace algorithms.

A recent study was developed by B. Bagasi et al. [37], It concerns the comparison between the SURF algorithm and the BRISK algorithm [38]. This is a study that deals with searching for images by content. It concerns the search for images by similarity in manuscript images without segmentation.

3. The Proposed Method

OCR optical character recognition software allows the conversion of characters in the image into the code (ASCII). These OCRs are only valid for printed text even though the recognition rate is not optimal for Arabic printed characters.

In this paper, we propose a method of searching the text in images of Arabic documents like that used in images of Latin manuscripts. In this context, due to the lack of methods for recognizing manuscripts, it is difficult to access their content in image form. Annotation or transcription techniques require considerable effort when entering text equivalent to the content of manuscript images. For this reason, these techniques can be automated by adding the Word Spotting method. Our goal is to facilitate the identification of the words in each scanned manuscript in the form of images. It is a question of characterizing by a signature each handwritten element (word, pseudo-word, letter). This signature must be unique for words that are identical. The objective is to eliminate redundancies in repeating elements.

3.1. Principles of the proposed method

The method of identifying and searching for Arabic handwritten elements is illustrated in schematic form by a block diagram. We have an input target image containing one page of a manuscript and two types of input queries for which the proposed method uses the text query or image query containing the word to be searched. In the case, of the text entry that is common, we used a database in XML format containing the features of the words of each manuscript processed. In case the input is in the form of an image, we use the direct extraction of the features of the request image. In other words, our method is based on text-image matching for identification of the contents of the manuscripts. In the following we deal with the case of comparison of input query images with a target image.

After the feature comparison operations, the results are provided in a final step as an output, thus locating the detected words. The following figure shows a block diagram of the proposed Word spotting method:

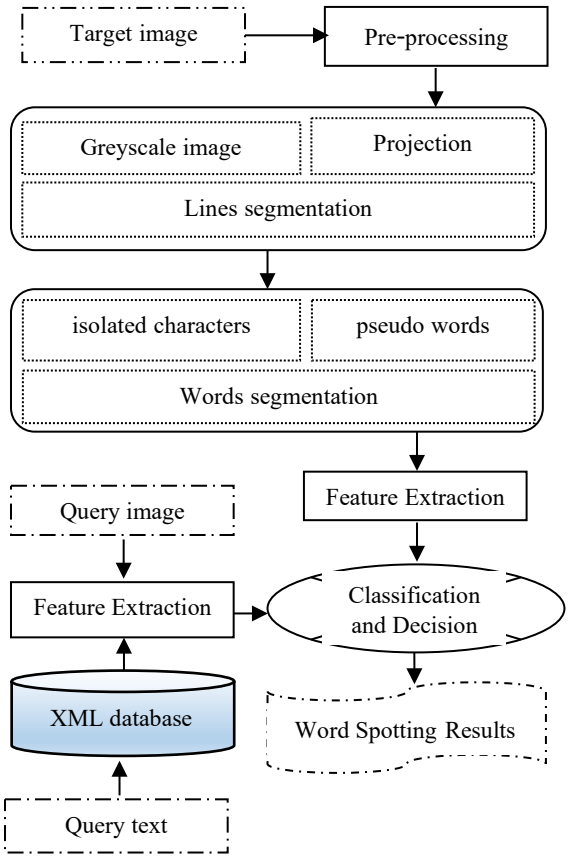


Figure 1: Synoptic diagram of the proposed Word Spotting method

- **Acquisition and pre-processing**

The first step is the acquisition of the manuscript in image form. Professional scanners make it possible to provide images with acceptable resolutions without damaging the manuscripts by the light emitted during scanning. Most often we obtain an image that contains two pages (left and right) in its raw state. We used as test images, manuscripts from the BNRM National Library in Rabat, Morocco. We also used images from a database of the Refaiya family [39] at the library of the University of Leipzig. The following figure shows the test image we used.



Figure 2: Original image [39]

After scanning, a series of pre-processing is required. Knowing that the scanned images are in their raw state in folders. Raw images take up a lot of space on storage disks. Image compression is necessary to reduce the storage capacity on the disk and speeds up processing time. Most of the professional scanners dedicated to manuscripts contain software and image processing algorithms. Other pre-processing must be applied to manuscript images such as: straightening, curvature correction, level spreading, contrast enhancement or detail enhancement, etc.



Figure 3: Left and right pages after pre-processing

In this article, we used the image on the right-hand side of the page. In most cases, Arabic manuscripts are written in different colours. We see in this example, that we have the colours of the red and black writing. The purpose of extracting text from the physical medium without losing important information. Colour segmentation of the image is an effective solution. It consists in selecting the pixel ranges in different colour layers. The most appropriate solution for our case is to segment each layer of the image separately and to assemble these layers with threshold values used. The HSV (Hue, Saturation, Value) space offers a simple way to detect thresholds. It allows easily distinguish intensity, saturation and hue.



Figure 4: Colour segmentation

- **Line segmentation**

The Word Spotting method requires a line segmentation [40] and subsequently a word segmentation to be able to locate the words to be searched in the image. In the first step, we will convert the colour image to a grayscale image whose purpose is to have a 1D signal. The projection algorithm [41] provides an effective way to facilitate line detection despite its sensitivity to noise. In this case, we use grayscale images instead of binary images because the projection of binary images [42] is very sensitive to line overlaps. It is widely used in the case of printed text because of the remarkable spacing between the lines. For our case, the projection of the grayscale image is represented by summing the intensity I at the coordinates of the image (x,y) to have the function of f(y) according to the following relation:

$$f(y) = \sum_{x=0}^{w} I(x, y)$$

The function f (y) represents the projected profile of image I. The calculation of the sum of pixels x along the y-axis is shown in Figure 5. Smoothing is essential to remove additional noise from the f (y) function. The convolution of a Gaussian with this function allows to filter the high frequency noise as in [41].

$$p(y) = f(y) * g(y, \sigma)$$

$$g(y, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{y^2}{2\sigma^2}}$$

The grayscale image projection profile provides several useful information about the text in the image. It allows to illustrate the various characteristics like the basic lines (minima), the heights of the lines and the background of the image (maxima). In this case,

the texture effect is not considered because we have added a pre-processing step as a colour segmentation. The maxima noted on the curve represent the spacing between the lines. The maxima (local) can be calculated by applying the derivative of the p(y) function.

$$\frac{dp(y)}{dy} = 0$$

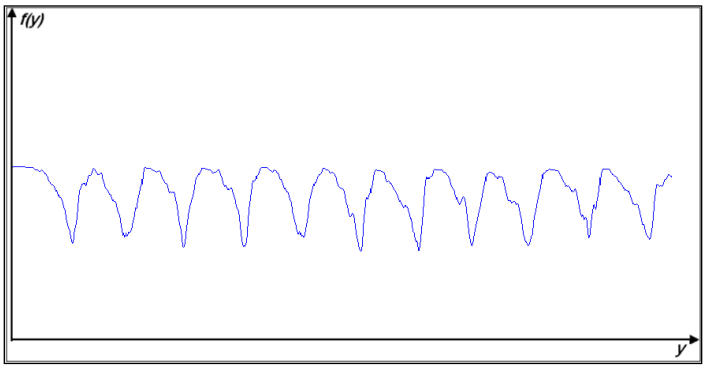


Figure 5: Line projection profile

The detection of local extrema facilitates the localization of the beginning and the end of lines. The following image shows the segmentation of the lines:



Figure 6: lines segmentation

- **Word segmentation**

The image in Figure 4, which is segmented in colour. It will be made binary by applying a global thresholding method. Among the thresholding methods, there is that of optimal Otsu thresholding [43] which we can apply thanks to the bimodal nature of the histogram of this type of image. As an effect, the resulting image is affected by binary noise. As a solution, binary morphological filtering reduces the effect of this generated noise. We have already identified the text lines in the previous step of line segmentation. Dilation by applying a morphological filter of the binary image along each line allows the merging of pseudo-words and characters. The projection technique at the level of each line provides words. Overlapping characters prevent segmentation by this projection technique. The application of labelling of related components at the level of each line of the binary images makes it possible to locate isolated characters, pseudo-words and words. This method is also sensitive to noise, but it is effective. Additional noise from small areas is not considered by our algorithms. The following figure shows an example of segmentation of handwritten elements by applying labelling of related components (words, pseudo-words and characters).



Figure 7: Words and elements segmentation results

- ***Feature extraction***

Arabic calligraphy has undergone several evolutions. It is considered an art of millennial history. The forms of Arabic letters follow refined rules. They change shape according to their position in the words (initial letter, median letter, final letter or isolated letter). Vowels can be designated by diacritical signs on words. Filtering of these diacritic signs is necessary to increase the quality of segmentation and recognition of words.

In most Arabic manuscripts, the writings follow uniform patterns according to the rules of the calligraphers. In many cases, these models are only the best-known fonts such as (Kufi, Andalus, Thuluth, Naskh, etc.). Words in manuscript images can be repeated several times with the same writing styles. The idea is to search for similar words in the same manuscript or in a collection of a calligrapher whose purpose is to use the occurrences of these words by our method.

Our objective is to choose such a method to extract the characteristics of words to be searched in manuscript images. We note that such a choice must be justified according to several constraints. In the literature, several approaches deal with recognition of forms. We opted for the use of invariable local detectors and descriptors to identify the forms of word writing. We justify the choice of these detectors by their acceptable response times and by their efficiency. On the other hand, we tested other methods, and could not obtain good results such as the theoretical signal model approach or the contour approach. We have the same character of the Arabic manuscript text that changes shape which makes the storyteller unstable. We opted for the approach based on local extrema because it uses points of interest recognized by their efficiency and by their optimised calculation times. In the literature, several approaches based on local extrema such as SURF descriptors offer a means of detecting these interest points.

In our case, the detection of words in images of Arabic manuscripts must take into consideration the styles of scripture. The approach using local extrema detection is based on variations in light intensity. However, in the case of small variations in light intensity, especially in manuscripts with almost identical intensities, the detection of word characteristics will be low. Another parameter to consider is the algorithmic cost and detector reliability. For example, the SIFT detector offers the interest points, with a higher computation time than the SURF detectors. That's why we use SURF's interest points to characterize the words in our Word Spotting method.

- ***Classification and matching***

The degree of similarity of two words concerns the comparison of the characteristics of these two words. In our case, the characteristics of the words are based on SURF interest points. Comparison between points of interest requires a choice of the appropriate method to optimize processing time. Interest points of are characterized by their properties and their associated descriptor vectors. We compare these interest points in two stages:
The first is to compare the trace sign of the Hessian matrix. This sign is from the Laplacian. It provides information on the meaning of blobs. Subsequently, two interest points to compare should have the same sign, if not different points.

The second is to compare the descriptor vectors of two interest points. Among the methods that can be used, the Euclidean distance or the Mahalanobis distance. The comparison is based on the calculation of distance and the correlation between these vectors.

4. Results and Discussion

Among the methods that are based on the Word Spotting technique we can mention the HADARA80P method [44] which uses a dataset of 80 pages available on the internet [45] including images in Tiff format at high resolution. The documents handled are 9th century Arabic manuscripts. The authors of this method use three types of requests: by string, by polygon and by image which includes 25 test requests. Compared to the HADARA80P results, our proposed method provides satisfactory results even for low resolution images. However, the HADARA80P method requires high-quality images of about 300 dpi, especially when it comes to keyword images. In this case, images exceeding 50MB capacity can make their broadcasts difficult on the internet.
In our case, we used a microcomputer whose features is: CPU i5 Quad Core with a frequency of 2.4 GHz, 3 MB cache, 4 GB RAM and operating system Windows10 64-bit. We did tests on hundreds of pages and as well as on the dataset of HADARA80P. With twenty images request, we have achieved 95.27%-word matching accuracy with a run time of up to 910 ms for high memory capacity images.

Our method is to characterize the Arabic manuscripts. It identifies handwritten elements such as words, pseudo-words, isolated characters. The manuscripts we process are in the form of 2D images. The physical medium used in most manuscripts has a uniform texture and produces small variations in light intensity. However, ink generates large variations in light intensity. Indeed, the SURF points of interest used by our method characterise the text instead of the texture. Small variations in light intensity have a negligible effect on the detection of SURF interest points in manuscript images.

4.1. Word Matching

Word matching goes through several stages. These steps include extracting the features of words as interest points. Each word is represented by a set of interest points. These points are invariable to changes in scale, geometric variations, rotations and brightness variations. These invariances are shown by tests applied to the different words that are considered areas with a set of interest points. Regarding the effect of the texture of the physical medium was shown by the queries images 1, 2 and 4 on the table. In this case, we show as well that the texture allows to provide a negligible number of points. In addition, the accuracy of word matching depends on the number of interest points detected on each word. Many interest points allow to refine the description of each search word. To have high accuracy when matching words, it is necessary to use high quality images without increasing the execution time when searching for words.

The table shows the number of points detected on each query image. For this purpose, the interest points of the query images are located by small blue disks on these images. About the number of matching points between each query image and the target image is shown in the last column of the table.

The query images are compared with the target image with several of points detected in the target image equal to 773. We have a maximum of match points equal to 423 of the query image number 4 in the target image. However, the query image number 1 gives a minimum number of match points equal to 89.

The following figure shows all the query images used in the previous table. The target image is the one in the middle of the figure.

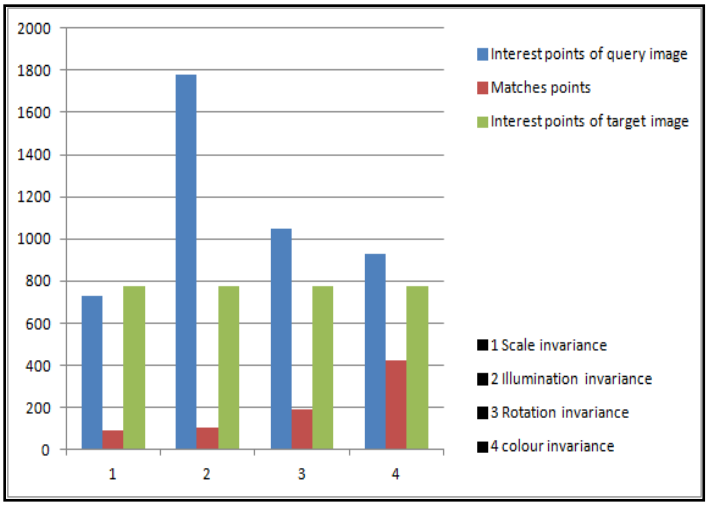
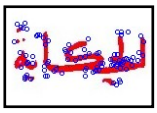
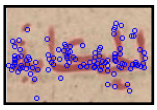
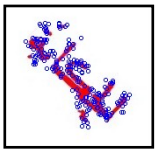
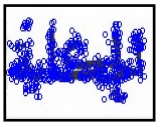


Figure 8: Histogram matching of interest points

Table 1: Experimental results of the matching points in q*uery images with a target image*

Query number	Invariance	Interest points	Query image	Matches points
1	Scale invariance	732		89
2	Illumination invariance	1782		107
3	Rotation invariance	1046		191
4	colour invariance	927		423

4.2. Application development

It is an application that is developed based on all the steps mentioned above (line segmentation, word and pseudo-word segmentation, feature extraction, classification). The extraction of the features is performed at the level of both query and target images. Word segmentation is an important tool to guide classification and identification. We used the SURF algorithm for the detection and description of features in the context of points of interest. This algorithm is known for its speed and robustness of its interest points in terms of invariance to scale changes, rotation, geometric variations and luminance. The following figure shows the interest points located in the target image and the query image on the left.

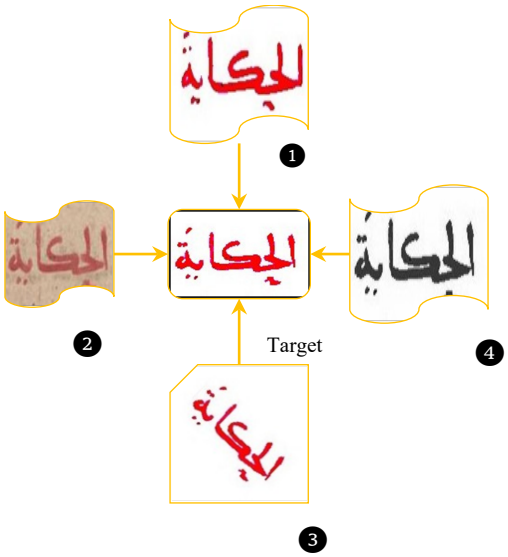


Figure 9: Four query images and the target image in the center



Figure 10: Interest points represented in the query images and in the target image

The objective is to develop a tool that allows access to the content of digitized Arabic manuscripts. The tools available for indexing handwritten documents in image form are very limited. In most cases, it is a search in databases describing the content of images with metadata. Our goal is to develop an application to identify text contained in images whose main interest is to integrate the Word Spotting technique for the search of occurrences of image word in handwritten documents. For this application, there are two possible cases:

- ✓ In a first case, the query image words are compared with those of the target image. Our method is based on interest points. However, an interest point belonging to a query image may belong to several words of the target image. Hence the need to segment the words of the target image and extract the features of each segmented word. Our interest is to search for all occurrences of a word image query in all pages of the same manuscript. The results of this research are acceptable despite some rejections due to segmentation problems in case of overlapping lines.
- ✓ In a second case, the query words can be text in ASCII format. This involves searching for the equivalent of text words in a database that stores the features of these words. Subsequently, the search for words in the target image is based on the features of the search words in the database.

The following figure shows the application we developed for the Word Spotting method. In the example, the comparison of the image query "history 'Al Hikayat' الحكاية" with a page image of an Arabic manuscript. The number of occurrences of this query image is equal to five times. As a perspective, we can expand this application to automate page annotations or help with transcription.

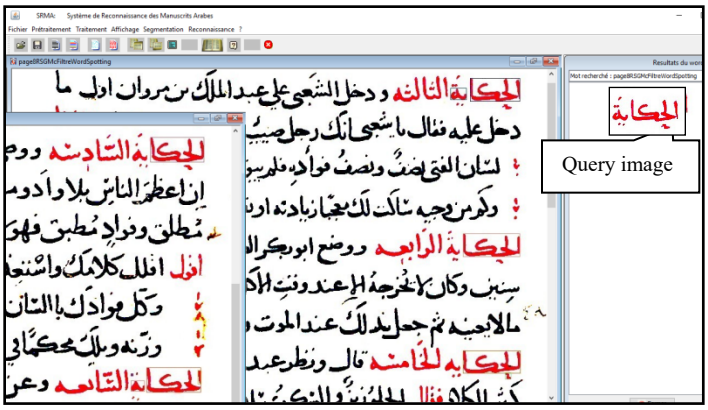


Figure 11: Word Spotting results

Conclusion

In this article, we have highlighted a method of searching for words in Arabic manuscripts. Most existing software (OCR) for Arabic text is only valid for printed text.
To this aim, we have proposed a word identification method based on the Word Spotting technique. The strength of this method lies in extracting the robust features of segmented words (or other elements: characters, pseudo-words). We use the invariant local detectors and descriptors of the SURF algorithm to extract the points of interest characterizing the words. The physical paper (or parchment) of the document has a negligible effect on the appearance of false points of interest belonging to the texture of the manuscript.

We have applied our method to hundreds of Arabic manuscripts from the National Library of Rabat and the Digital Library of Leipzig University. The test results of our method are excellent although some are very degraded. The limits of our method are in the case where the styles of writing and the fonts are varied, especially in the case the baldness of the lines where segregation is impossible.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] N. El makhfi, "Handwritten Arabic Word Spotting Using Speeded Up Robust Features Algorithm" in IEEE 5th International Conference on Optimization and Applications (ICOA), Kenitra, Morocco, 2019. https://doi.org/10.1109/ICOA.2019.8727692

[2] Bay H., T.Tuytelaars, L.V. Gool, "SURF : Speeded Up Robust Features", in 9th European Conference on Computer Vision, Graz Austria, 404-417, 2006. https://doi.org/10.1016/j.cviu.2007.09.014

[3] University of Leipzig. Project for the Cataloguing and Digitising of Islamic Manuscripts, www.islamic-manuscripts.net

[4] R. Manmatha, C. Han, and E. M. Riseman, "Word spotting: A new approach to indexing handwriting", in Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), San Francisco, CA, USA, 1996. https://doi.org/10.1109/CVPR.1996.517139

[5] T. Rath, R. Manmatha, and V. Lavrenko, "A search engine for historical manuscript images", in Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR), Sheffield, United Kingdom, 2004. https://doi.org/10.1145/1008992.1009056

[6] T. Rath and R. Manmatha, "Word spotting for historical documents", IJDAR, **9** (2), 13-152, 2007. http://dx.doi.org/10.1007/s10032-006-0035-8

[7] I. Yalniz and R. Manmatha, "An efficient framework for searching text in noisy documents", in DAS, Gold Cost, QLD, Australia, 2012. https://doi.org/10.1109/DAS.2012.18

[8] V. Frinken, A. Fischer, R. Manmatha, and H. Bunke, "A novel word spotting method based on recurrent neural networks", in IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), **34** (2), 211 – 224, 2012. https://doi.org/10.1109/TPAMI.2011.113

[9] A. Fischer, A. Keller, V. Frinken, and H. Bunke, "HMM-based word spotting in handwritten documents using subword models", in ICPR, Istanbul, Turkey, 2010. https://doi.org/10.1109/ICPR.2010.834

[10] F. Chen, L. Wilcox, and D. Bloomberg. "Word spotting in scanned images using hidden markov models". In Acoustics, Speech, and Signal Processing, ICASSP-93., IEEE International Conference on, vol **5**, 1–4, Minneapolis, MN, USA, 27-30 April 1993. https://doi.org/10.1109/ICASSP.1993.319732

[11] A. Fischer, A. Keller, V. Frinken, and H. Bunke, "Lexicon-free handwritten word spotting using character hmms", Pattern Recognition Letters, vol. 33, no. 7, pp. 934–942, 2012. https://doi.org/10.1016/j.patrec.2011.09.009

[12] D. Zhang and G. Lu, "Review of shape representation and description techniques", PR, **37**(1), 1–19, 2004. https://doi.org/10.1016/j.patcog.2003.07.008

[13] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection", CVPR, 2005. https://doi.org/10.1109/CVPR.2005.177

[14] J. Almazan, A. Gordo, A. Fornes, and E. Valveny, "Efficient Exemplar Word Spotting", in BMVC, 2012. [Online]. Available: http://www.bmva.org/bmvc/2012/BMVC/paper067/paper067.pdf

[15] S. Sudholt and G. A. Fink, "A Modified Isomap Approach to Manifold Learning in Word Spotting", in 37th German Conference on Pattern Recognition, ser. Lecture Notes in Computer Science, Aachen, Germany, 2015. https://link.springer.com/chapter/10.1007/978-3-319-24947-6_44

[16] J. Almazan, A. Gordo, A. Fornes, and E. Valveny, "Word Spotting and Recognition with Embedded Attributes" Transactions on Pattern Analysis and Machine Intelligence, **36** (12), pp. 2552–2566, 2014. https://doi.org/10.1109/TPAMI.2014.2339814

[17] L. Rothacker and G. A. Fink, "Segmentation-free query-by-string word spotting with bag-of-features HMMs", in International Conference on Document Analysis and Recognition, Tunis, Tunisia, 2015. https://doi.org/10.1109/ICDAR.2015.7333844

[18] M.Rusinol, D.Aldavert, R.Toledo, J.Llados,"Efficient segmentation-free keyword spotting in historical document collections" Pattern Recognition, **48** (2), 545–555, 2015. https://doi.org/10.1016/j.patcog.2014.08.021

[19] D. Aldavert, M. Rusinol, R. Toledo, and J. Llados, "Integrating Visual and Textual Cues for Query-by-String Word Spotting", in International Conference on Document Analysis and Recognition, 511–515. Washington, DC, USA, 2013. https://doi.org/10.1109/ICDAR.2013.108

[20] David. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints". **60** (2), 91–110, 2004. [Online]. Available: https://link.springer.com/article/10.1023/B:VISI.0000029664.99615.94

[21] David G. Lowe, "Object Recognition from Local Scale-Invariant Features". Proc. of the International Conference on Computer Vision, Kerkyra, Greece, 1999. https://doi.org/10.1109/ICCV.1999.790410

[22] Cordelia SCHMID, "Pairing of images by local gray-scale invariants, Application to the indexing of a base of objects", Ph.D thesis, National Polytechnic Institute of Grenoble, France, 1996.

[23] C. Harris and M. Stephens, "A combined corner and edge detector", Proceedings of the 4th Alvey Vision Conference. 147-151, 1988. [Online]. Available: http://www.bmva.org/bmvc/1988/avc-88-023.pdf

[24] Schmid, & R. Mohr, "Local grayvalue invariants for image retrieval", IEEE Transactions on Pattern Analysis and Machine Intelligence. **19** (5), 530–534, 1997. https://doi.org/10.1109/34.589215

[25] P. Montesinos, V. Gouet, and R. Deriche, "Differential invariants for color images", in IEEE 40th International Conference on Pattern Recognition, Brisbane, Queensland, Australia, 1998. https://doi.org/10.1109/ICPR.1998.711280

[26] Y. Dufournaud, Cordelia Schmid, Radu Horaud, "Matching Images with Different Resolutions", in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Hilton Head Island, SC, USA, 2000. https://doi.org/10.1109/CVPR.2000.855876

[27] K. Mikolajczyk, C. Schmid, "Indexing based on scale invariant interest points", in IEEE 8th International Conference on Computer Vision (ICCV) , Vancouver, BC, Canada, 2001. https://doi.org/10.1109/ICCV.2001.937561

[28] K. Mikolajczyk, "Detection of local features invariant to affine transformation", Ph.D thesis, institut national polytechnique, Grenoble, France, 2002.

[29] K. Rohr, "Recognizing corners by fitting parametric". International Journal of Computer Vision, **9**(3): 213-230, 1992. [Online]. Available: https://link.springer.com/article/10.1007/BF00133702

[30] k. Rohr, "Über die Modellierung und Identifikation charakteristischer Grauwertverläufe in Realweltbildern", In 12 DAGM-Symposium Mustererkennung, 217-224, 1990. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-84305-1_25

[31] R. Deriche & T. Blaska, "Recovering and characterizing image feature using an efficient model based approach". In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, New York, NY, USA, 1993. https://doi.org/10.1109/CVPR.1993.341079

[32] P. Brand & R. Mohr, "Accuracy in image measure". in Procedings of the SPIE Conference on Videometrics III, **2350**, 218-228, Boston, MA, United States, 1994. https://doi.org/10.1117/12.189134

[33] R. Deriche & G. Giraudon, "Accurate corner detection: an analytical study". In Proceedings of the 3rd International Conference on Computer Vision, Osaka, Japan, 1990. https://doi.org/10.1109/ICCV.1990.139495

[34] R. Deriche & G. Giraudon, "A computational approach for corner and vertex detection". International Journal of Computer Vision, **10**(2): 101-124, 1993. [Online]. Available: https://link.springer.com/article/10.1007/BF01420733

[35] T. lindeberg, "Feature detection with automatic scale selection", International Journal of Computer Vision, **30** (2), 79–116, 1998. [Online]. Available: https://link.springer.com/article/10.1023/A:1008045108935

[36] K. Mikolajczyk, & C. Schmid, "A performance evaluation of local descriptors", IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI, **27** (10), 1615 – 1630, 2005. [Online]. Available: https://doi.org/10.1109/TPAMI.2005.188

[37] B. Bagasi, L. A. Elrefaei, "Arabic Manuscript Content Based Image Retrieval: A Comparison between SURF and BRISK Local Features", International Journal of Computing and Digital Systems ISSN (2210-142X) Int. J. Com. Dig. Sys. **7** (6), 355-364, 2018. [Online]. Available: http://dx.doi.org/10.12785/ijcds/070604

[38] S. Leutenegger, M. Chili, and RY. Siegwart, "BRISK: Binary Robust invariant scalable keypoints" in 2011 IEEE International Conference. Computer Vision (ICCV), 2548-2555 , 2011, Barcelona, Spain https://doi.org/10.1109/ICCV.2011.6126542

[39] The Refaiya Family Library, Leipzig University Library. https://www.refaiya.uni-leipzig.de/content/index.xml?lang=en

[40] N. El makhfi, O. El bannay. "SCALE-SPACE APPROACH FOR CHARACTER SEGMENTATION IN SCANNED IMAGES OF ARABIC DOCUMENTS". Journal of Theoretical and Applied Information Technology, **94** (1), 2016 . [Online]. Available: http://www.jatit.org/volumes/Vol94No2/19Vol94No2.pdf

[41] R. Manmatha, N. Srimal, "Scale space technique for word segmentation in handwritten manuscripts", In 2nd International Conference on Scale Space Theory in Computer Vision (Scale-Space 99). **1682** (2), 22-33, Corfu, GRECE, 1999. [Online]. Available: http://ciir.cs.umass.edu/pubfiles/mm-27.pdf

[42] J. Ha, R. M. Haralick, and I. T. Phillips. "Document page decomposition by the bounding-box projection technique". In Proceedings of the Third International Conference on Document Analysis and Recognition (ICDAR), pages 1119 1122, Montreal, Quebec, Canada, 1995. [Online].Available: https://pdfs.semanticscholar.org/e896/99bf5a1593490d6d93699ca17f3c75434178.pdf

[43] N. Otsu, "A threshold selection method from greyscale histogram", in IEEE Transactions on Systems, Man, and Cybernetics. on SMC,.**9** (1), 62-66, 1979. https://doi.org/10.1109/TSMC.1979.4310076

[44] W. Pantke, M. Dennhardt, D. Fecker, V. Möargner, T. Fingscheidt, "An Historical Handwritten Arabic Dataset for Segmentation-Free Word Spotting -HADARA80P", in proceedings of the 14th International Conference on Frontiers in Handwriting Recognition (ICHFR),15-20, Heraklion, Greece, 2014. https://doi.org/10.1109/ICFHR.2014.11

[45] Technische Universität Braunschweig, HADARA80P Dataset. [Online]. Available: https://www.ifn.ing.tu-bs.de/research/data/HADARA80P

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 100-105 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Implement Wireless and Distributed Vibrator for Enhancing Physical Activity of Visually Impaired Children

Hung-Chi Chu[1], Fang-Lin Chao[*,2], Liza Lee[2], Pei-Yun Kao[2]

[1] *Department of Information and Communication Engineering, Chaoyang University of Technology, 436, Taiwan R.O.C.*

[2] *Department of Industrial Design,* [3]*Department of Early Childhood Development & Education, Chaoyang University of Technology, 436, Taiwan R.O.C.*

ARTICLE INFO

Article history:
Received: 27 October, 2019
Accepted: 08 January, 2020
Online: 15 January, 2020

Keywords:
Tactile sensation
Bodily activities
Visually impaired
Music
Wireless network

ABSTRACT

Learning and mental and physical counselling for visually impaired children are important. However, due to the lack of visual sense assistance, Physical Activity will gradually degenerate without proper enhancement, which may require the care of others for life. Inner fear and sensitive psychological factors hinder the development of living ability. Strengthening physical activity of visually impaired children improves their ability and reduces psychological obstacles, which is conducive to getting along well with others. The distributed vibration developed in this study can enhance children's physical exercise. Through music rhythm, children's psychological resistance is reduced, and their body movements are triggered by distributed vibration, which has an impact on the physical activity of children with visual impairment.

Seven kindergarten-age children were selected to participate in a 16-week tutoring. A score was recorded based on the assessment form, teaching record, and interviews with parents or students. The results showed that vibration and music are helpful in developing a young child's motor skills. After two sessions of tutoring, player became more interested in participating along with the music and tactile signals. Players were able to correctly identify the limbs that corresponded to the vibration and gesture with their hands and feet. Assistive technology had a positive effect on the physical activity of the visually impaired children.

1. Introduction

The daily life of the blind is not as active as ordinary people. The lack of visual sensory assistance results in problems such as slow movements and inability to respond immediately to the surrounding environment. Failure to integrate into normal life causes personality shrinkage, lack of contact with people, and decreased perception. With the advancement of technology, caring assistive technology has been applied to assist disadvantaged groups, such as: smart cane, route guidance, and voice/ image recognition. There were some facilities regarding to motor learning in visually impaired children [1,2,3]. Anthierens utilized a feel space navigation belt composed of 16 buzzers to distribute sensory cues for visually impaired kayakers [1]. Per Ellis presented sound therapy is effective for educating children who have multiple learning difficulties and may provide additional ways of communicating [4].

1.1. Haptic Devices & Teach

Haptic devices which interact with the user without disturbing them are most commonly used for drawing the attention of cell phone owners. Small direct current (DC) motors offer high performance and a quick response time for a low price. Piezoceramics convert electrical energy into mechanical vibrations. Anvik [5] had investigated the potential of rich tactile notifications with six vibrators.

Previous studies have also shown positive effects through airflow and tactile sensations [6] in promoting children's physical activity. Baker et al. developed a 'ring' to be worn by an orchestral conductor and a 'vibration matrix' through a two-dimension haptic signal through which the conductor could convey their intent to a visually impaired musician [7]. Low-cost vibrators had also been constructed to promote general distribution of the application to disadvantaged group [8].

* Fang-Lin Chao, Email: flin@cyut.edu.tw

1.2. Purpose

The aim of this study was to examine the effect of distributed vibration in enhancing the motor control of visually impaired children with disabilities. Both quantitative and qualitative methods were utilized to measure the validity of the intervention. Next, six second-grade elementary school students participated in an activity which paired vibration events with song segments. During the session, players were able to correctly identify the limbs that corresponded to the vibration and gesture with their hands and feet.

Design actuators and tactile devices supported the effectiveness of distributed vibration in enhancing the movement of children. The specific purposes of the study were:

(1) To design affordable equipment for enhancing exercise behaviors.

(2) To examine the effectiveness of the equipment of distributed vibrator and music activities in enhancing physical activity in children.

2. Assistive Device Design

2.1. Concepts and circuit implementation

The initial concept art of the vibrator application is shown in Figure 1. This concept art proposed tiny distributed vibrators to prompt movement at specific positions on the user. Tactile signals cue the user to respond with an action at an individual location.

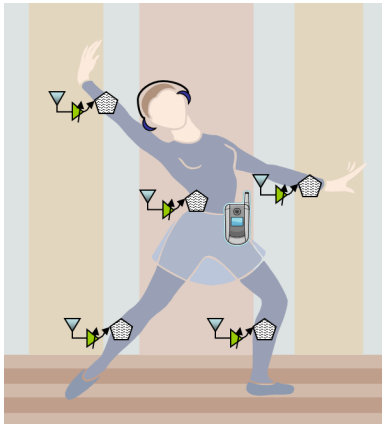


Figure 1: Concept art of distributed vibrators

For proper tactile cueing, the vibrators must contact the skin. In the early stage of the field test, we used elastic straps to attach the vibrator to the corresponding limb. Vibrators of the appropriate size received a signal through a wire; a computer controlled the on/off timing of each channel.

The prototype's circuit consisted of solid-state relay (SSR) components through a USB line. The control channel employed a Phidget Interface Kit 8/8/8 I/O board, which controlled the switching by preset sequence. An external DC power supply with solid-state relay was constructed (Figure 2) with a 360 ohm resistor in series with a 3.6V Zener diode to ensure the current amount.

Figure 3 shows the graphical user interface (GUI) of the control program. The interface was developed using Microsoft Visual Basic, which includes five functions. The first part is the tool bar to provide the connection, COM part setup, Record, Stop, Execute, Camera, and Exit functions. The second part is the melody loading, and the third part is the load/modify function of body movements, which can load preconfigured body movements. Users can choose specific vibrator through push button: LH (left hand), RH (right hand), LF (left foot), and RF (right foot) buttons to active movements. Finally, parts four and five show the body movement results.

After selecting the music, users pressed the top button to begin. The time and order of the button pressing were recorded in a text file for playback during the next test.

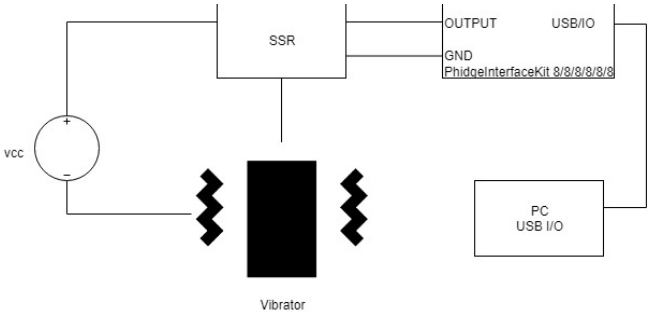


Figure 2: Hardware block diagram of the solid-state relay and vibrator

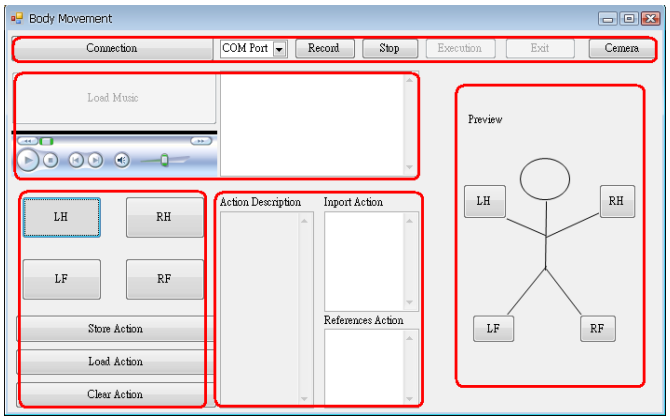


Figure 3: Editor screen layout of vibrators and control.

2.2. Wireless WSN implementation

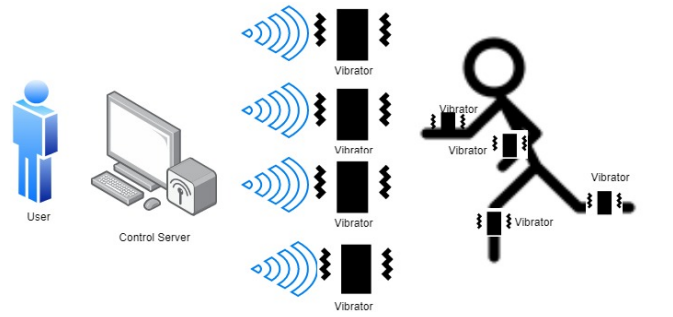
In addition, with the rapid evolution of telecommunications technology, assistive technology combined with microelectronics and information technology has benefits. Wireless sensor network is an infrastructure less, low cost, dynamic topology, application oriented design with wireless distributed nods. WSN designing become more complex for further processing [9]. Deployed sensors provided the required connectivity [10]. Sensor and the wireless implementation of the distributed vibrator are reviewed in this section.

A WSN emphasizes low-cost, low-speed, and ubiquitous communication between devices. The basic framework involves a 10-m communication range with a transfer rate of 250 kbps. The method significantly reduces line interference during movement. The system (Figure 4a) shows a central server using a ZigBee network interface, which is responsible for transmitting control signals and receiving messages from sensor nodes with the vibrator. The control command could turn on/off the vibrator with a specific time (Figure 4b). Each sensor module uses a 4.2 V lithium battery power. The vibrator is connected to a low-speed I/O pin on the sensor node to reduce the module's energy

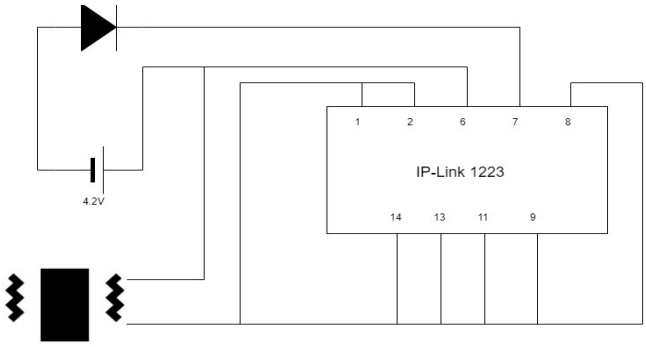
consumption — the hardware circuit prototype placed within a chassis made by 3D printing (Figure 5). The ZigBee interface is IP-Link5501, and the sensor node is IP-Link1223 [11].

Zigbee's Network Layer Protocol Data Unit (NPDU) shows in the Figure 6. An NPDU consists of a network header and a network payload. The Network header contains frame control of 2 octets, routing fields of 6 octets, and data payload NSDU (network service data unit) of variable length. Besides, the routing fields of 6 octets are composed of a destination address of 2 octets, a source address of 2 octets, a radius (broadcast radius) of 1 octet, and a sequence number of 1 octet. Among them: the destination address of 2 octets is set to an address of 0xFFFF to indicate broadcast information, thereby broadcasting the control signals of the four vibrators at the same time.

The sequence number of 1 octet is the broadcast sequence number (BCSN). The control field of the vibrator used in this research is to use the 1 byte in the original data payload to mark it as ID1, ID2, ID3, and ID4. Each ID represents a controlled vibrator, and its field length is 2 bits, in which the bit value code: 11 means intensity vibration, 10 means moderate vibration, 01 means low vibration, and 00 means stop. ID1 and ID2 represent the left, right-handed vibrator, ID3 and ID4 represent the left, right-foot vibrator.

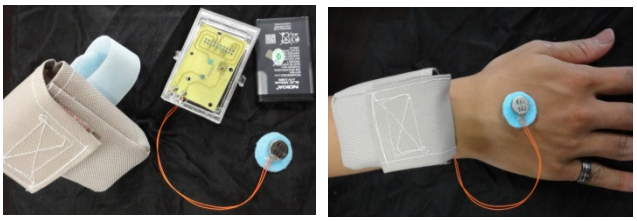


(a)



(b)

Figure 4: (a) Architecture of the assisted body movement system, (b) the circuit of vibration node.



(a) (b)

Figure 5: (a) Vibration node prototype with ZigBee Interface (IP-link 5501) and Sensor node: (IP-link 1223) (b) hand wearing.

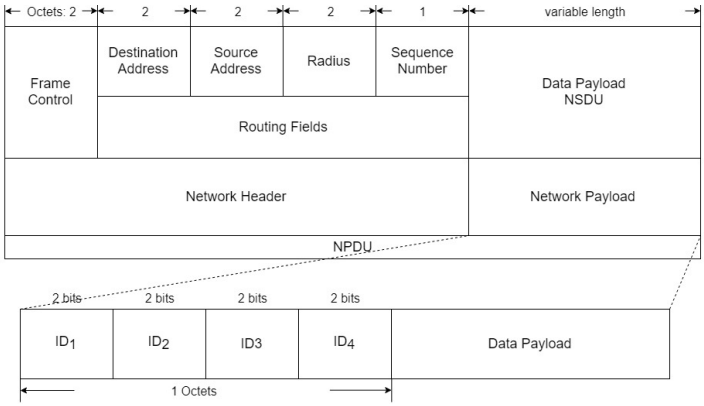


Figure 6: Network Layer Protocol data unit of Zigbee.

The increase of the volume of the module weakens the vibration amplitude, which affects the perception and response of the subject. The designer tried EVA and placed vibrator under the main circuit with an eccentric shaft. The main circuit is separated from the round flat vibration motor. The arc shape conforms to the curvature of the wrist, and the fabric and the devil felt are connected for convenience (Figure 7).

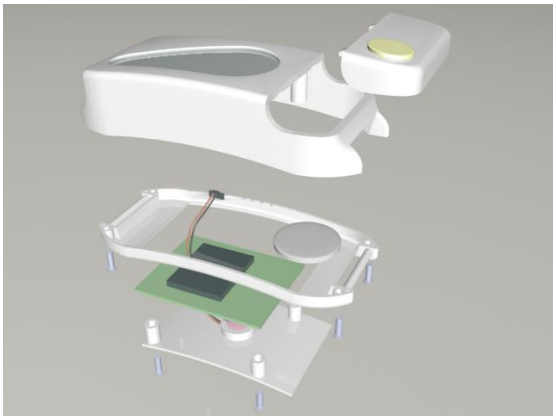


Figure 7: Module consists of an upper cover, chassis, and a battery cover. The chip is located on the lower chassis. There are power switch, mode setting button, and LED display switch.

2.3. Human Factor issues

Human factors design prompted ergonomic considerations for specific users. An interdisciplinary approach consist human factors, ergonomics, and work psychology in connection with design goals for cyber-physical product [12]. We kept the total weight of each vibrator under 130 grams through circuit miniaturization. The elastic straps followed the contours of the arm, wrist, or leg. Human factors were measured data. The maximum wrist circumference of the male wrist is about 19.8 cm and the average circumference of the adult ankle is about 25 centimeters. The initial design, featuring silicone material to help maintain skin contact with the backside-mounted vibrator, is shown in Figure 5.

The vibration amplitude/frequency need adjust for the specific user. Primary parameters of perception are: amplitude, frequency, timing, and location [13]. The skin is roughly sensitive to vibrations between 20 and 500 Hz. Vibration stimuli exceeds threshold will be detected. The threshold for the trunk is 4 microns at 200 Hz [14, 15]. High spatial resolution located on the hands, and face. Tactile displays must be setting between threshold and the maximum comfort level. Comfortable stimuli range 15-20 dB

above the absolute threshold. Amplitudes above 0.6-0.8 mm will result in a pain sensation. Design considerations reviewed considering: frequency, intensity, area of stimulation [16].

The stimulation responses of retarded children were analyzed in [17].Vibration sensation at low and high frequencies differed significantly. Sense thresholds lower frequencies (16 Hz and 31.5 Hz) were low. The relatively unchanged tactile frequency discrimination between vibration frequencies of 180 and 250 Hz were measured in [18].The movement of the device limited to a narrow bandwidth of frequencies below 30 Hz [19].

Devices for locating or stimulating reflex points in the body with alternating magnetic fields producing oscillating movement which preferably exerting perpendicular forces on the skin [20]. Children's availability and movement convenience were prioritized for location choosing. When setting the vibrator, avoid high frequency or intensity to prevent tactile shock (Figure 8).



Figure 8: Wired module and external power supply.

3. Methodology

Visually impaired children have different athletic abilities with age and innate physical conditions. We have arranged tests for two ethnic groups:

(1) Children with multiple impairments of varying degrees of severity, aged between 3-6 years old, have limited ability to autonomously move. Due to the combined effects of other obstacles, the ability to move is limited, and the personality is closed and passive. Physical exercise encourages them to try to show more variety or amplitude.

(2) Visually impaired children have normal ability - aged 9-12. They are amblyopic or totally blind, but their physical capabilities are close to the average person. Close observation reveals that students have limited ability to balance, with small and restrained movements. Usually, they are accompanied by each other and passively do the activities within the allowable range under the teacher's request. Students may have obesity due to less activity. Physical exercise encourages them to demonstrate amplitude rhythm and interaction.

Based on the purpose of the study, the specific research questions asked were:

(1) Could multiple vibrators and music activities assist a young child with visually impaired development of motor skills?

(2) Could multiple vibrators enhance a visually impaired child's range of movement?

3.1. Wired test

The wired module is first introduced, it had limitation of connected wire but the facility cost is low and easy to implement. Teachers can give exercise tips in a small area. Two phases of testing were arranged to understand the adaptation to the vibrator for different ability: Phase- I involved group (1) 3-6 years-old with multiple disable, and Phase- II anticipate group (2) 9-12 years-old with normal ability.

Quantitative analysis utilized measuring validity based on assessment scales, semi-structured reports, and video recording. The therapeutic procedures included the establishment of goals, assessment, and activity evaluation. Some teaching styles were chosen based on participants' learning responses. The study took place at a school in central Taichung, Taiwan. Seven visually impaired participants were selected. Three participants exhibited severe physical disability, and the other four children exhibited moderate physical disability. The Phase-I participants were given 40-minute tutoring session once per week for 16 weeks (Figure 9).



Figure 9: Field practice during Phase-I tutoring

The quantified results of a semi-structured musical activities observation form were compared with related background documents (reports from the parents and teaching logs). The forms were completed by a physical therapist in order to gather physical movement data which was scored by three observers.

3.2. Wireless test

The wireless module lifts the distance limitation of the vibration connection. In addition to prompting in a small range, you can also provide motion prompts in a large area covered by radio waves. The wireless module test also carried out to understand the adaptation to the vibrator for different ability group. The approaches were similar to wired test except utilized the wireless modules.

4. Results and Discussion

4.1. Results of wired test

Phase- I Observation

Figure 10 shows the participants' physical movements during the test. After attending the program, all participants exhibited significantly improved physical mobility. The three participants

with severe symptoms were able to move their heads and hands slightly. The other four participants with moderate symptoms were able to walk and run stably and change direction. The tutoring sessions had a positive result (Figure 11).

Week 1 to Week 4: At the beginning of the research, none of the participants had the motivation to perform the movements. When the instructor put the distributed vibration equipment on the participants' limbs, and did not know what to do. The four participants with mild to moderate symptoms tried to pull the attached accessory. The frequency of the vibration was high and fast, consequently the participator could not follow it accordingly. A preset guidance protocol was required

Week 5 to Week 11 (same experiment setup): After the team adjusted the frequency of the vibration signal, most of the children did not reject the vibration and began to move to the music. Those with mild and moderate symptoms were able to move when participator felt the vibration and heard music. Sometimes participator would tell the instructor "felt the vibration on skin". The researchers found that the fuse of the attached accessory and cables was too long; as a result, the vibration signal did not transmit properly sometimes. Therefore, the team improved the cable routing and electrical contact.

Week 12 to Week 16: All the participants made progress at different levels (Figure 11). The participants with severe symptoms showed facial expressions of happiness. The participants with less severe symptoms were able to respond to the vibration and music correctly. The performance is slowly increasing by the time not right after using this equipment. For the mental state of the visually impaired with multiple obstacles, they couldn't accept changes from outside quickly. People with multiple disabilities have weak movements and expressions. After getting used to tactile sensations and being able to respond appropriately, additional channel of communication prompts them to participate more.



Figure 10: Participants' physical movement with (a) severe and (b) moderate symptoms

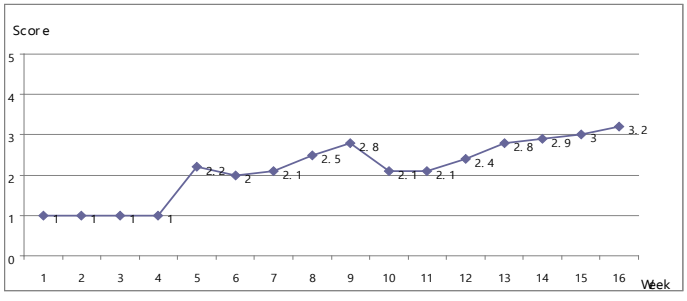


Figure 11: Observers' assessment of participants' physical movement

Phase- II Observation

Vibrators were placed on the upper and lower body (subjects' hands, wrists, and ankles). Six elementary school students participated in a five-minute session where vibration events were paired with segments of a song. The vibrators were turned on, along with the popular song, "Sorry," by Super Junier. Subject C01 (Figure 12a) appeared initially frightened and covered her ears. After the teacher's guidance, C01 was able to complete the directed actions.



(a) (b) (c)

Figure 12: (a) C01's reaction, (b) C02, (c) C04/05situation

At first, Subject C02 (Figure 12b) found the experience strange and a little frightening. However, after guidance, C02 became more relaxed and willing to jump in place. Subject C03 swung her body around and occasionally moved her wrist in small motions. After some instruction, Subject C03 was able to wave her left arm. Subject C04 was more lively, swinging his body rhythmically and creating his own dance moves. He shook his body during the interlude and appeared to be having fun. Subject C05 (Figure 12c) could correctly identify which vibrator was vibrating and raise and lower his foot, but his rhythm was not as good as C04's. Both participants corresponded to the vibrations and made appropriate gestures with their hands and feet. Exercises make them more aware of their relative positions.

4.2. Results of Wireless usage

Ten special children are invited to perform small-scale sports, including children with low vision, blind children, and slow intellectual development. After instructing the child to equip the child with a vibrator, observe the response with the music and understand the differences in the way of using it.



Figure 13: Children spontaneously tap the instrument

Children spontaneously tap the instrument according to the vibration instructions. Special children feel keen, and some children want to take off the vibrator because their hands are not comfortable. A child with a slow intellectual development has an absolute sensation of music show outstanding performance. Two

student laugh and very focus on the indicator and tap the instrument. Some children complete the test through the guide of the accompanier. (Figure 13)

5. Conclusions

Through music activities with children of varying ages, the findings of the study were as follows: The participants' capacity for physical movement improved significantly. We also observed an improvement in social validity from the teachers and participants. The tactile design was able to assist the development of the participants' physical skill. The wireless module lifts the distance limitation of the vibration connection. In addition to prompting in a small range, you can also provide motion prompts in a large area covered by the radio waves.

Vibrator devices reinforce active participation for children with severe disabilities. Students exhibited high willingness to withstand new stimuli. Therefore, it is necessary to adapt a teaching strategy to the characteristics of the students. Wireless technology enables remote instructions to be given to assist users in interacting with the people and environment. This is not something that can be achieved by purely hardware design. The effectiveness of the equipment design also relies on developing the teaching strategy and events.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported in part by the National Science Council, Taiwan, ROC, under grant NSC 101-2221-E-324-028, and 99-2221-E-324-026-MY2.

References

[1] Anthierens, Cédric, Didier Groux, and Vincent Hugel, "Sensory navigation guide for visually impaired sea kayakers" Journal of Field Robotics, **35**(5), 732-747, 2018.
[2] Zuzanna Maćkowiak, Wieslaw Osiński,"The effect of sensorimotor training on the postural stability of visually impaired women over 50 years of age" Journal of Women & Aging, 27(1), 68-80, 2015. DOI: 10.1080/08952841.2014.928140
[3] Malik, S. et al., "Orientation and mobility training in special education curriculum for social adjustment problems of visually impaired children in Pakistan" International Journal of Instruction, 11(2), 185-202, 2018.
[4] Phil Ellis, "Incidental music: A case study in the development of sound therapy," British Journal of Music Education, 12(1), 59-70, 1995.
[5] Cecilie H. Anvik, "Embodied spaces in the making: visually impaired people, bodies and surroundings" Scandinavian Journal of Disability Research, 11(2), 145-157, 2009. DOI: 10.1080/15017410902830710
[6] Schwellnus, H., et al. "Using movement-to-music technology for play with children with special needs" OT Now, July (2002).
[7] Baker, David, Ann Fomukong-Boden, and Sian Edwards, "'Don't follow them, look at me!': Contemplating a haptic digital prototype to bridge the conductor and visually impaired performer" Music Education Research 21(3), 295-314, 2019.
[8] Bischof, A. et al. "Exploring the playfulness of tools for co-designing smart connected devices: a case study with blind and visually impaired students" Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion,. ACM, 2016.
[9] Jangra, Ajay. "Wireless sensor network (WSN): Architectural design issues and challenges" 2010.
[10] Mohamed, Shaimaa M., Haitham S. Hamza, and Iman Aly Saroit, "Coverage in mobile wireless sensor networks (M-WSN): A survey" Computer Communications, 110, 133-150, 2017.
[11] IP-Link 1223, Helicomm Inc., http://www.helicomm.com/showproducts.asp?ID=144
[12] Stern, Hendrik, and Till Becker, "Development of a model for the integration of human factors in cyber-physical production systems" Procedia Manufacturing, 9, 151-158, 2017.
[13] Van Erp, J. B., "Guidelines for the use of vibro-tactile displays in human computer interaction" In Proceedings of eurohaptics, Vol. 2002, 18-22, 2002.
[14] Verrillo, R.T., "Effect of contactor area on the vibrotactile threshold" Journal of Acoustical Society of America, 35 (12), 1962-1966, 1963.
[15] Sherrick, C.A. & Cholewiak, R.W., Cutaneous sensitivity. In Boff, K.R., Kaufman, L. & Thomas, J.P.:Handbook of perception and human performance (pp. 12-1 –12-57), New York: John Wiley and sons, 1986.
[16] Verrillo, Ronald T. "Psychophysics of vibrotactile stimulation." The Journal of the Acoustical Society of America 77.1 (1985): 225-232.
[17] Pujol, Kelley Kruse. "The effect of vibrotactile stimulation, instrumentation, and precomposed melodies on physiological and behavioral responses of profoundly retarded children and adults" Journal of Music Therapy, 31(3), 186-205,1994.
[18] Harada, Noriaki, and Michael J. Griffin. "Factors influencing vibration sense thresholds used to assess occupational exposures to hand transmitted vibration" Occupational and Environmental Medicine, 48(3), 185-192, 1991.
[19] Wall, Steven A., and William Harwin. "A high bandwidth interface for haptic human computer interaction." Mechatronics 11.4 (2001): 371-387.
[20] Herman, Goldfarb. "Massaging garment with vibrators located in back and chest sections" U.S. Patent No. 3,310,050. 21 Mar. 1967.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 104-110 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

New Color Image Encryption for Medical Images Based on Three Dimensional Generalized Chaotic Cat Map and Combined Cellular Automata

Un Sook Choi[1], Sung Jin Cho[2,*], Sung Won Kang[2]

[1]*Department of Information and Communications Engineering, Tongmyong University, Busan, South Korea*

[2]*Department of Applied Mathematics, Pukyong National University, Busan, South Korea*

ARTICLE INFO

Article history:
Received: 13 December, 2019
Accepted: 05 February, 2020
Online: 09 March, 2020

Keywords:
Image Encryption
Chaotic Cat Map
Combined Cellular Automata
Medical Image Encryption
PRNG
Group Cellular Automata

ABSTRACT

Medical images are transmitted via the Internet or the hospital intranet which include many important information about the patient's personal information. Medical image encryption is a technology that can effectively protect the information contained in these medical images. In this paper, we give a secure and trusty Combined Cellular Automata (CoCA) based medical image encryption algorithm. CoCA is made up of a nonlinear CA and a linear 90/150 maximum length CA, and three dimensional generalized chaotic map. The proposed algorithm consists of two phases. The first phase encryption process that changes pixel values of a given image using two CoCAs. The second phase is to change the position of each pixel in the image encrypted using the three dimensional generalized chaotic cat map that can change with effect the position of pixels not only in two dimensions horizontally and vertically but also in color plane three colored channels simultaneously. We show the stability and high reliability of the given color medical image cryptosystem through detailed analysis and various statistical experimental tests.

1. Introduction

With the development of internet communication technology, the amount of transmission of various types of multimedia data including images is increasing. This trend is also true for medical data. Medical diagnostic images are transmitted over wired and wireless networks. These medical images are transmitted via the Internet or the hospital intranet which include many important information about the patient's personal information. However, current hospital intranet lacks reliable security equipment. In addition, on the Internet, not only malicious manipulation of information by malicious attackers, but also leakage of information on unauthorized third parties is very serious. Medical image encryption is an effective technique that can protect not only medical images but also important personal information of patients from these various attacks.

Some existing traditional encryption techniques, such as International Data Encryption Algorithm (IDE), AES and DES are well suited for text data security, but not for bulk data encryption such as digital images. This is because images have inherent properties with large data sizes, high redundancy, and strong imagery correlation.

Since chaos function is ergodic and sensitive to initial conditions and parameters, it is a powerful function that can effectively encrypt various types of data. These chaotic functions are easy to implement with microprocessors and personal computers [1]. Therefore, the chaos map-based encryption algorithm is suitable for mass multimedia data encryption because it can design a fast encryption system at low cost.

Recently, methods of medical image encryption based on chaotic function have been studied by many researchers [2-10]. Mao et al. proposed a cryptographic algorithm that not only maintains a high level of security but also speeds up image encryption using a map that extends the Baker map in three dimensions [2]. In [3], Dai and Wang proposed logistic maps and Chebyshev maps based a medical image encryption. They overcame the problem that logistic maps-based cryptographic systems are inadequate for medical image cryptography because of their low key space and low security. Wang et al. designed a color image cryptographic system that used the chaotic function to simultaneously encrypt the three colored components of the given image and allow these components to affect each other [4]. Fu et al. proposed an alternative bit shuffling algorithm in which pixel value mixing effects can be contributed by the replacement and the permutation process. That chaotic based algorithm reached

*Sung Jin Cho, e-mail: sjcho@pknu.ac.kr

existing cryptographic security levels with fewer rounds of performance [5]. Zhang and Luo proposed chaotic CA based an image encryption algorithm [6]. The algorithm proposed by them was used a two dimensional logistic function to encrypt images and then applies two dimensional CA on the initial encrypted image. In [7], Nandi et al. noted that the unique and powerful nature of CA is the ease of hardware implementation and the large number of rules that apply to the CA, making it impossible for an attacker to find a combination of rules that is the key to extracting the original information. So they proposed a CA-based cryptographic algorithm to substitute pixel values of original image. However, this method needs to be supplemented because it is weak to noise and data loss that may occur for image transmission.

Jeong et al. gave an encryption algorithm for medical image. They used a maximum length one-dimensional CA with complemented rules and a modified chaotic cat map. They changed pixel values of the given image using a nonlinear key sequence generated by a one-dimensional CA with complemented rules. And then they changed pixel positions of the image with pixel value changed using the modified chaotic map. However, the modified chaotic maps they designed have an unmaintained weakness of the image size required by the chaotic cat map [8]. Therefore, there is a need to overcome the weaknesses of the modified two-dimensional chaos function in [8] and to efficiently encrypt color images. In order to enhance security, a PRNG that can generate a nonlinear key sequence is required.

In this paper, we design CoCA for medical image encryption. CoCA is made up of a nonlinear CA and a linear 90/150 maximum length CA. And we give a secure and trusty encryption scheme based on CoCA and three dimensional generalized chaotic map. The proposed algorithm consists of two phases. The first phase encryption process that changes pixel values of a given image using two CoCAs. The second phase is to change the position of each pixel in the image encrypted using the three dimensional generalized chaotic cat map that can change with effect the position of pixels not only in two dimensions horizontally and vertically but also in color plane three colored channels simultaneously. We show the stability and high reliability of the given color medical image cryptosystem through detailed analysis and various statistical experimental tests.

2. Preliminaries

2.1. Cellular Automata

Von Neumann [11] introduced CA at first and Wolfram [12] developed CA. CA is very important mathematically for modelling complex behavior. CA is composed of an array of basic memory called cells, and its state is updated to the following state depending on the state of neighboring cells of radius 1 including itself. These CAs are classified into one-dimensional CA, two-dimensional CA, and the like according to the arrangement of the cells. In addition, according to the number of cells that affect when one cell of the CA is updated to the following state, it is classified as 3-neighbor CA, 5-neighbor CA, 9-neighbor CA. And it is classified into NBCA, PBCA, and IBCA according to boundary condition of both ends of CA. The CA used in this paper is an NBCA. The CA used in this paper is the simplest CA that is the 1-D 2-state 3-neighbor NBCA.

The state transition function f, which decides the following state of the cell, is a mapping in which the domain is {000, 001, 010, ... 111} and the airspace is {0,1}. f is expressed as follows.

$$w_j^{t+1} = f(w_{j-1}^t, w_j^t, w_{j+1}^t) \quad (1)$$

, where w_j^t is the jth cell state of CA at the time t. As shown in Table 1, the CA transition rule is a decimal representation of the numerical value of the following state determined according to the state of the current neighbor. For example, if the following state function value for {111, 110, 101, ... , 010, 000} is '01010010', the converted value is 82, so the given transition rule is 82. Table 1 shows the transition rule numbers of the state transition function according to the following state of the cell with respect to the state of the current neighbor cells.

Table 1: The Transition Rule Number according to the Following State of the Cell

Nbd state	111	110	101	100	011	010	001	000	RN.
Following state	0	0	0	1	1	1	1	0	30
	0	1	0	1	1	0	0	1	90
	1	0	0	0	1	1	0	1	141
	1	0	0	1	0	1	1	0	150
	1	0	1	0	1	1	0	0	172

When the state transition function is expressed as a Boolean function, a rule consisting only of XOR logic is called a linear rule, and a rule consisting of XOR logic and XNOR logic is called a complemented rule. And rules that include OR and AND logic are called nonlinear rules. In Table 2, rules 30, 141, and 172 are nonlinear rules, and rules 90 and 150 are linear rules. A CA having a linear rule applied to all cells of a CA is called a linear CA, and a CA to which a nonlinear rule is applied is called a nonlinear CA. The linear CA can express a function for obtaining the following state as a matrix, which is called the state transition matrix T_n.

Table 2: The Boolean functions for Rules 30, 90, 141, 150 and 172

RN	Boolean functions
30	$w_i^{t+1} = w_{i-1}^t \oplus (w_i^t + w_{i+1}^t)$
90	$w_i^{t+1} = w_{i-1}^t \oplus w_{i+1}^t$
141	$w_i^{t+1} = w_i^t \cdot w_{i+1}^t \oplus w_{i-1}^t \cdot w_{i+1}^t \oplus w_{i-1}^t \oplus w_{i+1}^t \oplus 1$
150	$w_i^{t+1} = w_{i-1}^t \oplus w_i^t \oplus w_{i+1}^t$
172	$w_i^{t+1} = w_i^t + w_{i-1}^t \cdot w_{i+1}^t$

From now on we use a CA using only the transition rules 90 and 150 which is called 90/150 CA. T_n of the 90/150 CA is as following [13, 14] :

$$T_n = (t_{ij})_{n\times n} = \begin{cases} 1\ , & i-j=1, i-j=-1 \\ d_i\ , & i=j \\ 0\ , & otherwise \end{cases} \quad (2)$$

The linear CA $\boldsymbol{C}$ can express a function for obtaining the following state as a matrix T_n. We call T_n the state transition matrix of $\boldsymbol{C}$. The CA applicable to the cryptographic algorithm must be reversible and decryptable. This reversible CA must be unique to the previous state for a given state. Therefore, the inverse of T must exist. This means that the determinant of T must not be zero. If W^t is the current state of $\boldsymbol{C}$ and W^{t+1} is the following state of W^t, then W^{t+1} can be obtained by T_n and W^t as follows [15] :

$$W^{t+1} = T_n W^t \quad (2)$$

By (2), the state W^{t+k} satisfies the following equation :

$$W^{t+k} = T^k W^t \quad (3)$$

An n-cell reversible CA $\boldsymbol{C}$ is referred to as Maximum Length CA (MLCA) when the minimum p satisfying $T_n^p = I_n$ for T_n of $\boldsymbol{C}$ is $2^n - 1$, where I_n is the $n \times n$ identity matrix. For example, a four-cell CA with transition rules <90,150, 90,150> is MLCA with period 15.

2.2. Three Dimensional Generalized Chaotic Map

For a metric space X with metric d, a function $F: X \to X$ that satisfies the following is said to be sensitive to initial conditions:

if there is a $\delta > 0$ such that for every $x \in X$ and $\varepsilon > 0$ there exists $y \in X$ with $d(x, y) < \varepsilon$ and $\mathrm{d}\left(F^k(x), F^k(y)\right) > \delta$ for $k \in \mathrm{N}$ [16].

Chaos is associated with randomness, disorder, and entropy. Arnold Cat Map is a simple function proposed by Russian mathematician Arnold that can randomly transform the composition of image pixels. This transformation causes the image to be very complicated when this transformation is repeated by changing the position of the pixels without changing the size of the image. This function is represented by a two-dimensional matrix

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \quad (4)$$

and has been applied to various algorithms to protect images. Chen et al. generalized (4) using parameters as shown in (5) [17].

$$\begin{pmatrix} 1 & a \\ b & 1+ab \end{pmatrix} \quad (5)$$

Jeong et al. designed an encryption algorithm for medical color image using the map modified (5) [8]. However, (5) only changes the pixel position of the gray image because it can change the pixel position of the two-dimensional image. Therefore, in order to effectively protect color images, methods of extending from the 2-D chaotic cat map to a 3-D chaotic cat map have been proposed [17, 18]. A 3-D chaotic cat map is gotten by the following operations using the 3×3 identity matrix:

$$c_j \leftarrow a_{ij} \cdot c_i + c_j, r_j \leftarrow b_{ij} \cdot r_i + r_j \quad (6)$$

, where $a_{ij}, b_{ij} \in N,\ i = 1,2,\ j = 1,2,3\ and\ i \neq j$. Using (6), we obtain L_1 for $i = 1, j = 2, 3$.

$$L_1 = \begin{pmatrix} 1 & a_{12} & a_{13} \\ b_{12} & a_{12}b_{12}+1 & a_{13}b_{12} \\ b_{13} & a_{12}b_{13} & a_{13}b_{13}+1 \end{pmatrix} \quad (7)$$

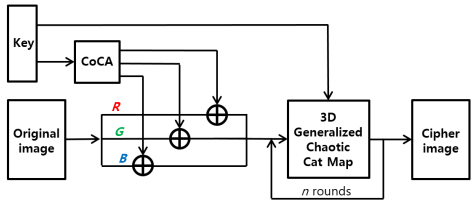
Similarly using (6) we get L_2 for $i = 2, j = 1, 3.$, when. The three dimensional generalized chaotic cat map is as follows:

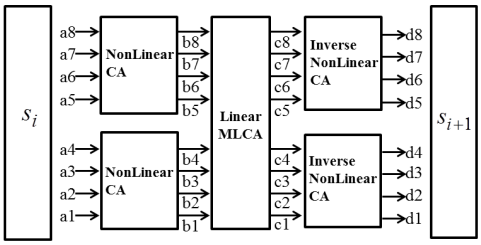
$$L = L_2 L_1 \quad (8)$$

Since $\det(L_i) = 1$ for each $L_i (i = 1, 2)$, $\det(L) = 1$.

3. Proposed Algorithm

Now we design an algorithm for encrypting medical images. Figure 1 shows a block diagram of the proposed algorithm. The medical image encryption algorithm proposed in this paper consists of two phases. The proposed algorithm consists of alternate and shuffling phases, and the pixel shuffling phase repeats n rounds to increase the security level of the encryption system.





3.1. Substitution Phase

To generate an effective key image we design an effective PRNG. This PRNG is constructed as shown in Figure 2 by the combination of nonlinear reversible CA C_1 and linear maximum length CA C_2. This PRNG is called CoCA. When S_0, a shared key, is the initial value of CoCA, the following state S_1 is $S_1 = C_1^{-1}C_2C_1(S_0)$. And $S_2 = C_1^{-1}C_2^2C_1(S_0)$. Therefore, S_n obtained by repeatedly applying CoCA n times to S_0 satisfies $S_n = C_1^{-1}C_2^nC_1(S_0)$. Therefore, even if C_1 is not an MLCA, if C_2 is an MLCA, the CoCA becomes a nonlinear MLCA. The algorithm proposed in this paper generates key images using two CoCAs G_1 and G_2. The (i, j) pixel of the key image generated by G_1 and G_2 is represented $K_{i,j}$. $K_{i,j}$ can be expressed as (9):

$$K_{i,j} = G_2^i\left(G_1^j(S_0)\right)\ (1 \le i, j \le 256) \quad (9)$$

r / c	1	2	3	4	5	·····	253	254	255	256
1	96	109	57	61	47	·····	203	208	112	34
2	217	50	91	121	166	·····	141	111	11	111
3	182	105	233	82	239	·····	63	29	4	139
4	52	105	43	68	84	·····	43	194	66	98
5	155	155	89	75	176	·····	146	13	79	230
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
253	194	66	98	207	0	·····	36	38	6	41
254	109	187	47	148	63	·····	75	54	203	22
255	22	160	207	134	9	·····	25	18	246	178
256	146	13	79	230	224	·····	82	239	244	210

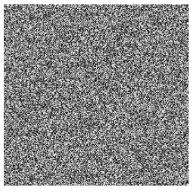


Figure 3 shows the procedure of generating a key image from two CoCAs. This process divides the shared key value 8 bits S_0 in half and performs state transitions through 4 cell nonlinear reversible CA C_1, and then combines them into 8 bits again. The combined 8-bit values are once again transitioned using 8-cell linear MLCA C_2. This transitioned result is then divided in half and converted to the final state by $C_1^{-1}$the inverse of C_1 to produce pixel values. First, each pixel of the first row of the key image is a

key sequence generated using G_1, and pixels of each column of the key image are generated by G_2 using the pixel value generated by G_1 as the initial value.

3.2. Shuffling Phase

The result of the XOR combining of the key image generated by two CoCAs and the original image is very effective because it results in the randomly changing the pixel value of the given image. However, if the encrypted image is damaged by unexpected noise or deliberate deletion attacks in the process of transmitting the digital image, the damaged part remains completely lost when the encrypted image is decrypted. Therefore, in order to design a cryptographic algorithm that is resistant to such attacks, in this paper, the final cryptographic image is generated by mixing the pixel positions of the image encrypted by the key image using a generalized 3-D chaotic cat map. Each pixel in the image has a gray level value from 1 to 255 for the R, G, and B components. Therefore, the total number of bits required to represent a color image of size M × M is M × M × 3 × 8. Each pixel position of a color image can be represented by (row, column, color component). In the image shuffling phase, we use the L in (8) to change not only the position of the rows and columns of pixels but also the positions of the components of R, G, and B using a generalized 3-D chaotic cat map. Mathematically, the pixel shuffling phase is represented by (10).

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \\ z_{n+1} \end{pmatrix} = L \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} \tag{10}$$

In (10), if each row of L is $L(r_i)(i = 1,2,3))$, then $x_{n+1} = L(r_1)(x_n, y_n, z_n)^T mod\ M$, $y_{n+1} = L(r_2)(x_n, y_n, z_n)^T mod\ M$, $z_{n+1} = L(r_3)(x_n, y_n, z_n)^T mod\ 3$, where M is the multiple of 3.

3.3. Color Medical Image Encryption Algorithm

Table 3 shows a color medical image encryption algorithm Col_Medi_Encryp_Algorithm. This algorithm is a color image encryption algorithm that merges the encryption phases proposed in sections 3.1 and 3.2.

Table 3: Col_Medi_Encryp_Algorithm

Function 1 : CoCA state transition
Def CoCA*(Image, rule_table.nonlinear, rule_table.MLCA, Rule_table.Inverse)*
For (*i, j, k*) in *Image_Index*(*row, column, color*)
a = *rule_table.nonlinear*(*Image[i, j, k]* //16)
b = *rule_table.nonlinear*(*Image*[*i, j, k*]%16)
Image[*i, j, k*] = a*16+b
number = *key.MLCA*
For (*i, j, k*) in *Image_Index*(*row, column, color*)
Key_Image[i, j, k] = ((*number*∧*rule_table.MLCA*) ⊕ (*number*//2) ⊕ (*number**2)) mod 256
Image = *Image* ⊕ *Key_Image* mod 256
For (*i, j, k*) in *Image_Index*(*row, column, color*)
a = *rule_table.Inverse*(*Image*[*i, j, k*] // 16)
b = *rule_table.Inverse*(*Image*[*i, j, k*] % 16)
CoCAed_Image[*i, j, k*] = a*16 + b
Return *CoCAed_Image*

Function 2 : 3D generalized chaotic cat mapping
Def 3D_Chaotic_cat_map(*Image, cat_map_matrix, times*)
Construct (*Index_list(row,column,color) of Image*)
For i from 1 to *times*:
cat_mapped_Index_list = *matrix_vector_multiply*(*cat_map_matrix, Index_list*(*row, column, color*) of *Image*)
For (*i, j, k*) in *Image_Index(row, column, color)*
Cat_mapping_Image[*i, j, k*] = *Image*[*cat_mapped_Index_list*([*i, j, k*])]
Return *Cat_mapping_Image*

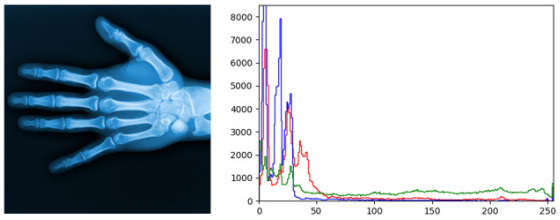
Algorithm : Col_Medi_Encryp_Algorithm
Input : *Image, rule_table.nonlinear, rule_table.MLCA, Rule_table.Inverse, cat_map_matrix, times*
Output : *Encrypted_Image*
Encrypted_Image = CoCA(*Image*, *rule_table.nonlinear, rule_table.MLCA, key.MLCA, rule_table.Inverse*)
Image = 3D_Chaotic_cat_map(*Image, cat_map_matrix, times*)

3.4. Key Scheming

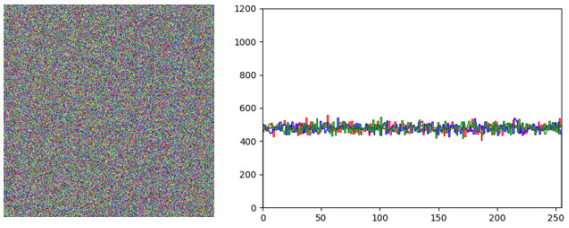
The variables that can be used as keys are: (1) Transition rules and initial values, which are the components of the two CoCAs require for the generation of a key image in the substitution phase. (2) Parameters of the proposed chaotic cat map used in the shuffling phase of mixing pixel positions.

4. Safety analysis of the proposed system

By the proposed cryptosystem we present the performance results in this section. Since medical images often have a high color bias, it must be shown whether the results are suitable for such images. Medical X-ray images were adopted as basic images suitable for such safety analysis.



(a) Original image and histogram



(b) Encrypted image and histogram

Figure 4: The histograms of the original image and the encrypted image

Table 4: Entropy and Comparison Analysis of Methods

Methods	Entropy
Proposed Algorithm	7.99837
Wang et al. [19]	7.99763
Wu et al. [20]	7.99722

4.1. Histogram Analysis

Figure 4- (a) shows the original medical image and histogram used to perform encryption. Most of the given medical images occupy black, and the original image actually has a very high bias at pixel value zero.

Figure 4- (b) shows the image and histogram that encrypted the original image using the proposed encryption system. Although the original medical image is a high deflection image, the pixel values of the encrypted image are very evenly distributed. This means that the distribution of colors is spread evenly according to the result of the encryption system. This proves to be safe for traditional statistical methods.

4.2. Information Entropy Analysis

A cryptographic image with high uncertainty means it is safe from accidental information attacks. In other words, the encrypted image must have a high degree of uncertainty with respect to information analysis. High uncertainty can be expressed through information entropy. Information entropy is as shown in (11).

$$H(m) = \sum_{i=0}^{2^n-1} p(m_i)\, log_2(1/p(m_i)) \quad (11)$$

Where m is the information source and $p(m)$ is the probability for finding the original value for each bit. Table 4 shows the information entropy of the encrypted image obtained by the proposed encryption system. In addition, the information entropy of cryptographic images obtained through several other cryptographic systems is also shown. In an 8-bit RGB format image, the closer to 8, the higher the uncertainty. According to Table 4, it is shown that the proposed cryptosystem, like the previously proposed cryptosystems, provides high uncertainty in the cryptographic image.

4.3. Correlation Coefficient Analysis

Correlation coefficients between pixel values indicate how distributed the image is in a particular color. The correlation coefficient is expressed as in (12) below:

$$\rho_{xy} = cov(x,y)/\sqrt{D(x)D(y)} \quad (12)$$

where x and y represent the values of two different positions in the RGB color space. In the correlation coefficient analysis, two positions are examined by randomly extracting two adjacent pixels horizontally, vertically, and diagonally. We examined random 2500 pairs of pixel values for each color channel of the image for correlation coefficient analysis. After repeating this test 400 times, the mean of all the extracted correlation coefficients was examined.

Table 5: Correlation Coefficients of an original image and an encrypted image

Methods	Horizontal	Vertical	Diagonal
Original image	0.9785	0.9730	0.9543
Proposed Algorithm	0.0019	0.0048	-0.0048
Zhang et al. [21]	0.0035	0.0037	0.0023

We provide the results of the correlation coefficient analysis obtained in the above-mentioned method in Table 5. Correlation coefficient analysis results are for the original image, the encrypted image using the proposed encryption, and the encrypted image based on previous studies. The value obtained according to (12) is distributed from -1 to 1. The closer to 1, the higher the correlation coefficient, which means that adjacent pixels have almost the same value. The correlation coefficients of the original image are close to 1 in either direction. This means that pixel values in adjacent locations almost always have a similar color. Closer to zero means that there is almost no correlation. The correlation coefficient of the encrypted image using the proposed encryption system is very close to zero in all directions.

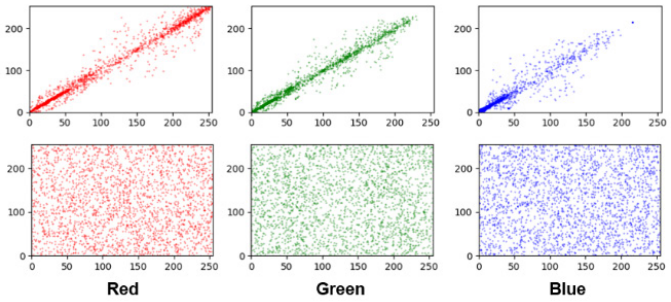


Figure 5: Scatter plots of horizontal direction of each color channels of the plain image and cipher image

The encrypted image indicates that it has different values at any adjacent positions, and it is difficult to find out what the original image is by color analysis. Figure 5 shows a scatter plot in the horizontal direction for each color channel in the normal and password images. It can be seen from the figure that the proposed encrypted image has little correlation between adjacent pixels.

4.4. Sensitive Analysis

An adversary can make a differential attack by analyzing different encrypted images obtained from a single key. Therefore, the slightest change of the image in the cryptosystem should cause a huge change as a whole. Analysis of these changes is called sensitivity analysis, and there are two well-known methods: The number of changing pixel rate (NPCR)

and the unified averaged changed intensity (UACI). NPCR and UACI for analyzing the proposed encryption algorithm are defined as the following (13) and (14):

$$NPCR_{RGB} = \frac{\sum_{i,j} D_{RGB}(i,j)}{W \times H} \times 100(\%) \quad (13)$$

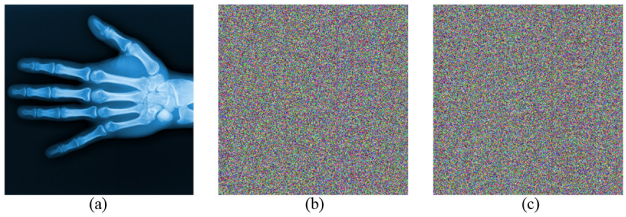
$$UACI_{RGB} = \frac{\sum_{i,j} |C_{RGB}(i,j) - C'_{RGB}(i,j)|}{W \times H \times 255} \times 100(\%) \quad (14)$$

where $W \times H$ represent the size of the plain image. In (14), $C(i,j)$ denotes a difference between pixel values of the ith rows and the jth columns of an image, and $C_{RGB}(i,j)$ and $C'_{RGB}(\mathrm{i,j})$ are of an original image and an encrypted image, respectively. In (13), $D_{RGB}(i,j) = 0$, if $C_{RGB}(i,j) = C'_{RGB}(\mathrm{i,j})$, and $D_{RGB}(i,j) = 1$, if $C_{RGB}(i,j) \neq C'_{RGB}(\mathrm{i,j})$.

Table 6: NPCR and UACI of encrypted images of different Images due to a small difference

Color	NPCR (%)	UACI (%)
R	99.6347	33.6082
G	99.6404	33.6748
B	99.61445	33.5853

NPCR and UACI of medical cryptographic color images through the proposed cryptographic system are provided in Table 6. The NPCR of over 99% indicates that the microscopic image changes almost entirely. In cryptography, about 33% of the UACI indicates that the distribution of differences in pixel values is even. Figure 6 shows encrypted images with different 1 bit of original image data.



(a) (b) (c)

Figure 6: Encrypted images with different 1 bit of original image data

4.5. Key Space Analysis

The number of available keys determines the size of the key space. To create a key image, we need two CoCAs consisting of a nonlinear CA and a 90/150 MLCA. In 8-cell CA, the number of rules is the number of cases applicable to each cell, which is 2^8. the number of rules that can be used for two CoCAs is $2^{8\times16} \cdot 2^{8\times16} = 2^{256}$. The number of keys required for the initial value is $2^{8\times3} = 2^{24}$. We can control the variables a and b for the pixel shuffling process of the 3D cat map, and the number of variables is N^{12} where N is the size of one direction of the image to be shuffled. If $N = 256$, the key space size is $2^{256+24+136} = 2^{416}$ in the proposed encryption scheme. Therefore, the proposed medical cryptography algorithm has a sufficiently large key space that the information eavesdropper cannot launch brute force attacks.

4.6. Restore for data corruption

Data stored in the database may be damaged by abnormal external interference. Image encryption algorithms must be able to resist these abnormalities. Experiments are conducted to test the resistance to noise and data loss. Figure 7 shows that even when abnormal data corruption occurs in an encrypted image, the original image is sufficiently resilient to be identified.

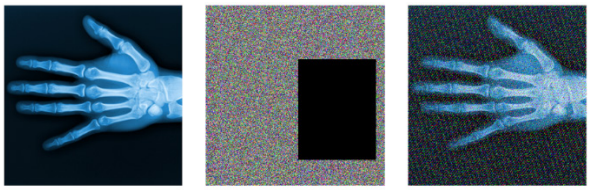


Figure 7: Restoration of damaged cipher image

5. Conclusion

In this paper, we proposed a safe and reliable color medical image algorithm based on CoCAs and three dimensional generalized chaotic cat map. The proposed algorithm had a very large key space and improved the nonlinearity of the key stream generated by using CoCAs. In addition, in the pixel shuffling phase of the encrypted image, the pixel positions of the image were changed together with the R, G, and B planes simultaneously, and thus we implemented a cryptosystem having a very strong resistance to noise and data loss. The detailed analysis of the encrypted images obtained by the proposed algorithm showed the high security and stability of the new medical color image encryption system. In the future, we will study high-speed image encryption suitable for encrypting large-capacity images and study various CA-based PRNG designs.

Acknowledgment

This paper is the revised and expanded version of a paper entitled “Color Medical Image Encryption using 3D Chaotic Cat Map and NCA” presented at 2019 10th IFIP International Conference on New Technologies, Mobility and Security (NTMS), Canary island, Spain, 24-26 June 2019. And this research was supported by the Tongmyong University Research Grants 2018(2018A046).

References

[1] X. Wang, L. Liu and Y. Zhang, “A novel chaotic block image encryption algorithm based on dynamic random growth technique,” Opt. Laser Eng., **66**, 10-18, 2015 https://doi.org/10.1016/j.optlaseng.2014.08.005

[2] Y.B. Mao, G,R, Chen, S.G. Lian, “A novel fast image encryption scheme based on 3D chaotic baker maps,” Int. J. Bifurcation Chaos, **14**(10), 3613-3624, 2004. https://doi.org/10.1142/S021812740401151X

[3] Y. Dai, X. Wang, “Medical image encryption based on a composition of logistic maps and Chebyshev maps", in 2012 IEEE International Conference on Information and Automation, Shenyang, China, 210-214, 2012. http://doi.org/10.1109/ICInfA.2012.6246810

[4] X. Wang, L. Teng, X. Qin, “A novel color image encryption algorithm based on chaos,” Signal Process., **92**(4), 1101-1108, 2012. https://doi.org/10.1016/j.sigpro.2011.10.023

[5] C. Fu, W. Meng, Y. Zhan, Z. Zhu, F.C.M. Lau, C.K. Tse, H. Mae, “An efficient and secure medical image protection scheme based on chaotic maps,” Comput. Biol. Med., **43**(8), 1000-1010, 2013. https://doi.org/10.1016/j.compbiomed.2013.05.005

[6] S. Zhang and H. Luo, “The Research of Image Encryption Algorithm Based on Chaos Cellular Automata,” J. Multimedia, **7**(1), 66-73, 2012. https://doi.org/10.4304/jmm.7.1.66-73

[7] S. Nandi, S. Chakraborty, S. Roy, W.B.A. Karaa, S. Nath, N. Dey, “1-D Group Cellular Automata Based Image Encryption Technique,” in 2014 IEEE International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), Kanyakumari, India, 521-526, 2014. https://doi.org/10.1109/ICCICCT.2014.6993017

[8] H.S. Jeong, K.C. Park, S.J. Cho, S.T. Kim, “Color medical image encryption using two-dimensional chaotic map and C-MLCA,” in 2018 10th International Conference on Ubiquitous and Future Networks (ICUFN), Prague, Czech Republic, 801-804, 2018. https://doi.org/10.1109/ICUFN.2018.8437025

[9] R.Enayatifar, H.J. Sadaei, A.H. Abdullah, M. Lee and I.F. Isnin, “A novel chaotic based image encryption using a hybrid model of deoxyribonucleic acid and cellular automata,” Opt. Laser Eng., **71(8)**, 33-41, 2015. https://doi.org/10.1016/j.optlaseng.2015.03.007

[10] Y. Zhou, L. Bao, C.L.P. Chen, “A new 1D chaotic system for image encryption,” Signal Process., **97**, 172-182, 2014. https://doi.org/10.1016/j.sigpro.2013.10.034

[11] J.V. Neumann, Theory of Self Reproducing Automata, University of Illinois, Urbana, 1966.

[12] S. Wolfram, “Computation theory of cellular automata”, COMMUN MATH PHYS, **96**(1), 15-57, 1984. https://doi.org/10.1201/9780429494093-4

[13] S.J. Cho, U.S. Choi, H.D. Kim, Y.H. Hwang, J.G. Kim and S.H. Heo, “New synthesis of one-dimensional 90/150 linear hybrid group cellular automata,” IEEE Trans. Comput-Aided Design Integr. Circuits Syst., **26**(9), 1720–1724, 2007. https://doi.org/10.1109/TCAD.2007.895784

[14] U.S. Choi, S.J. Cho, H.D. Kim, J.G. Kim, “Analysis of 90/150 CA corresponding to the power of irreducible polynomials,” J. Cell. Autom., 14(5-6), 417-433, 2019.

[15] P.P. Chaudhuri, D.R. Chowdhury, S. Nandi and S. Chatterjee, Additive cellular automata, Theory and applications, vol. 1, Los Alamitos; California; IEEE Computer Society Press, 1997.

[16] B. Hasselblatt, A. Katok, A First Course in Dynamics: With a Panorama of Recent Developments, Cambridge University Press, 2003.

[17] G. Chen, Y. Mao and C.K. Chui, “A symmetric image encryption scheme based on 3D chaotic cat maps,” Chaos Solitons Fractals, 21(3), 749-761, 2004. https://doi.org/10.1016/j.chaos.2003.12.022

[18] U.S. Choi, S.J. Cho, J.G. Kim, S.W. Kang, H.D. Kim, S.T. Kim, “Color Image Encryption Based on PC-MLCA and 3-D Chaotic Cat Map,” in 2019

IEEE 4th International Conference on Computer and Communication Systems (ICCCS), Singapore, 2019. http://doi.org/10.1109/CCOMS.2019.8821691

[19] X. Wang, Y. Zhao, H. Zhang, K. Guo, "A novel color image encryption scheme using alternate chaotic mapping structure," Opt. Laser Eng., **82**, 79-86, 2016. https://doi.org/10.1016/j.optlaseng.2015.12.006

[20] X. Wu, H. Kan, J. Kurths, "A new color image encryption scheme based on DNA sequences and multiple improved 1D chaotic maps," Appl. Soft Comput., **37**, 24-39, 2015. https://doi.org/10.1016/j.asoc.2015.08.008

[21] L. Zhang, Z. Zhu, B. Yang, W. Liu, H. Zhu, M. Zou, "Cryptanalysis and improvement of an efficient and secure medical image protection scheme," Math. Probl. Eng., **2015**(2), 1-11, 2015. https://doi.org/10.1155/2015/913476

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 106-111 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Comparative Study of Semantic and Keyword Based Search Engines

Bzar Khidir Hussan *

Department of Information System Engineering, Erbil Polytechnic University, Erbil- Kirkuk Road, Hadi Chawshli Street- Kurdistan Region- Iraq

ARTICLE INFO

Article history:
Received: 30 September, 2019
Accepted: 28 December, 2019
Online: 15 January, 2020

Keywords:
Conventional Search Engines
Syntactic Search Engines
Semantic Web
Intelligent Engines
Semantic Search Engines

ABSTRACT

Day by day, the data on the web becomes very huge which makes it difficult to find relevant information. Search engines are one of the successful factors that can retrieve information from the Web. The process of seeking information by search engines helps users find information on the internet, however it is not an easy task to find the exact information from this massive data available on the Web. Semantic Web technology has an ability to focus on metadata rather than syntax, which made the semantic search engines to search for the meaning of keywords instead of the keyword syntax. Consequently, an effective role of performance in conventional search engines can be achieved by rising the accuracy of information returned by a search query. In this paper, a survey for syntactic-based search engines and semantic-based search engines are studied, a comprehensive comparison between the two is presented, finally, their technologies are compared and discussed.

1. Introduction

The term Web is one of the most important technologies that allow users to access huge and different information through different locations in the world [1]. This information which is stored on servers is typically unstructured or - semi structured data [2,3]. The data on the Web is tremendously increasing which leads to many obstacles such as difficulty of finding relevant information or discovering exact knowledge on the Web [4].

The first generations of Web search engines, such as AltaVista, were indexing the contents of Web pages. The second generation of search engines, for instance Google, were considering the links to/from a Web page as a method of determining relevance. Both generations were mostly syntactic, which means they were depending on the keywords as text in their queries. Searching for interested information in most of the current search engines will result of retrieving hundreds of thousands retrievals while most of them are not relevant to what is meant to be found [5,6]. Hence, to solve such difficulties, Semantic search is intended to be updated so that it depends on the meaning of the keywords instead of the context [7].

Semantic Search Engines (SSE) use Semantic Web (SW) technology in their systems which make them intelligent to retrieve related information on the Web [8]. The SW aims at providing information as formal, well defined meanings, compatible, sharable knowledge base, and can be processed by machines [3]. Ontology acts an important role in the SW technology as it's famous as of the backbone of the SW structure, and is the vital element of SW infrastructure [7]. Web Ontology Language (OWL) and Resource Description Framework-Schema (RDFS) are the recommendations of the World Wide Web Consortium (W3C) for data representation models so as to deliver foundations for the ontology descriptions [4]. Ontologies provide distinct descriptions in their information, as a result, they are used in numerous fields and applications since its knowledge representation is understandable and processable by software agents and systems [8]. Ontology is a collection of semantically related concepts built on a limited number of predefined relations and terms of a domain. These terms and concepts can be represented visually so as to ease the representation for both syntactic and semantic data [9]. In Web, once abstract data is distributed across several knowledge bases, ontologies are the solely resolution as commonplaces to interpret the mutual senses of the domain key terms. Hence, significant concerns seek the development of ontologies [10].

Transforming syntactic search engines to semantic ones is not an easy task, since in the later one search results rely on the meaning of the query keywords, henceforth the search engines have to understand the keywords semantically in order to

* Bzar Khidir Hussan, Bzar.hussan@epu.edu.iq

retrieve relevant information. To do so, a new layer could be added to the so called syntactic search engines [11]

This paper presents the Web and SW technologies along with the Keyword and Semantic search engines. Keyword Search Engines (KSE)s are presented in section 2, section 3 reveals SSEs. In section 4, technologies of SSEs are presented, and section 5 exposes the most common SSEs. In section 6, a comprehensive study to the literature is given and a comparison of the works is presented. Section 7, gives a discussion panel for the studied systems. Finally, the suggestions and conclusions finalizing the research.

2. Keyword Search Engines

Conventional Search Engines are very helpful in finding information on the internet and getting results within some time, but they suffer from the fact that they do not know the meaning of the terms and expression used in the web pages and the relationship between them [12,13]. Surveys show that users who seek for information on the web do not find accurate results in the first set of URLs returned, because of increasing size of links on the web pages. Sometimes one word has several meanings and several words have the same meaning, in that case if a user wants to search for a particular word then it may produce confusion and user will not get what he wanted to search [2,14].

Search Engine Optimization (SEO) are used to find and search alternative search terms that people use with search engines while looking for similar subjects [15]. SEOs search for more keywords, which are used to achieve better rankings in search engines. Once a different keyword is found, they expand on it to find similar keywords for that keyword [16]. Keyword suggestion tools usually aid the process of finding similar keywords such as in [17] where substitutional keywords are the suggestions for the query. There are many techniques used in keyword search engines such as identify the core of the keyword, research related search terms, create a list of main terms and long-tail keywords, use the Google AdWords keyword planner…etc. [18].

3. Semantic Search Engines

The Semantic Search Engines (SSE)s are the intelligent engines that search for keywords depending on their meaning [19]. In addition, they guarantee the results those are related to the meaning of searched keywords. SSE use ontologies so that they achieve the meaningful retrievals and get a high accuracy result. [12]. SW considers as a Web 3.0, or an extension to the current web which represent information in order to link information in the web as a form of HTML, OWL and RDF files [5]. The SSEs are distinguished to have several types of relational links among verity types of resources instead of the single relation of resources. There are many examples of SSEs such as Hakia, DuckDuckGo, Swoogle… etc. The methods to store information within the SW technology are able to answer complex queries given to a search engine [17][12].

3.1. Technologies of Semantic Search Engines

The SSEs are essentially based on some technologies or methods which effectively achieve the SSEs [20]. These technologies are sometimes called the SW layers which include applying some tools such as inference engines, rule languages, annotation tools… etc. One of the main technologies used in the SSEs are the ontologies which can be offered within the form of RDF, RDFs, and OWL [21]. These technologies are used in the structure of SSE which briefly can be described below:

3.2. Unicode and URI

Defines as the base level of the SW technology which is used for identification and determination of the resource location. The Unicode is used to standardize the letters of machines, while the URI is used for identifying each resource by unique name [22].

3.3. Extensible Markup Language (XML)

A subset of Standard Generalized Markup Language (SGML) and machine readable represented by markup language. This language is widely used in the web for some reasons such as it is simple, flexible text format and its structure used to describe data. XML meets the challenges of E-business and electronic publishing as well. In addition, it has a very important role to exchange between different kind of data on the web [23].

3.4. Resource Description Framework (RDF)

It is representing the primary layer used in SW. RDF is very important to represent data which can be processed by machines [17]. The method that is used to identify and provide the relationship among the resources called graph model. The best simple model language which is used in this layer is RDF Schema which is used to create relations and descriptions of resources [24].

3.5. Ontology Vocabulary

The ontologies are used to describe data of the SW and improve methods to give uniform way to make easy communications among different parts of resources and be understandable by each other. Ontology is the method that can provide common grammar and vocabulary of data that are published, specifically the description of semantic data which represent by ontology [18].

3.6. Logic, Proof and Trust

Are the last layers of the SW cake which follow the ontology layer. These layers are used to check and solve consistency problems and the trustworthiness of SW. In addition, redundancy of concept duplicate data [25].

3.7. Common Semantic Search Engines

The SW technology has grown a new generation of the web by using some new methods to search about the best results related to searcher intent. There are many engines which depend on the semantic approach [17]. In the following sections, the most common SSEs are presented.

3.8. Hakia

It considered one of the common SSEs which provide results that are relevant to the concept of words instead of main

keywords. In other side, it's not just depending on the keywords but use the concepts of entire phrases such as questions or sentences. The one of the most important characteristics of Hakia is its capability to provide the results depending on equivalent concept such as "cure=treat" or "cons=disadvantages" [12]. The results of the search are divided into classes such as Web, News, and Video etc.; also, they can be divided according to the date or relevance. The technology which is used in Hakia is the OntoSem technology, which is a linguistic database [19]. The words here are categorized into different senses by depending on the QDEX (Query indexing technique), which considered as an infrastructure to index the data by Semantic Rank algorithm. This algorithm use the ontology and fuzzy logic to search about all possible requests [18].

3.9. Kngine

The Kngine divides the search results into two types: documents or images. This engine searches for the information related to the search term which means search about the concepts of the words [12]. It is very intelligent engine because its retrievals related to natural of question. For example, search about the city, the expected results will be related to the city lactation, events, weather and history [20].

3.10. DuckDuckGo

It has many of features that distinguish of other SSEs. When a keyword is being searched the results will have many related retrievals, meaning that, it provides different answers for the searcher, the searcher can choose the answer that is related to his intent [17]. For example, when we search about the term Apple, the engine will provide many of answer such as fruit, computer, bank, etc. [26]. The DuckDuckGo is distinguished among other types of semantic search engines by dealing with all users the same results when search the same term. Also, it deals with many other websites including Wikipedia, Bing and yahoo [27].

Table 1: Main Differences Between Keyword and Semantic Search Engines

No	Semantic search Engines	Keyword Search Engines
1	An approach to achieve a related and accurate information to user queries [12].	Considered as traditional engines which return the results depending on context of queries [18].
2	It depends on the stop words and punctuation marks. These marks effects on the search results [16].	It does not depend on the stop words and punctuation marks and the results are not accurate [21].
3	The OWL and FDF languages are the base languages used to creating web pages [19].	The HTML, XML and CSS languages are the base languages used for creating the web pages [12].
4	Seeks to provide the accuracy in returned information by understanding the meaning of keywords related to what the seeker desires [20].	It is searching exactly depending on words in the website which determined by searcher [22].
5	Tries to access to relations among main words by using the ontology [22].	Tries to expand the query by using keywords instead of using the methodology [25].

Table 2: Pros and Cons of Semantic Search Engines

No	SSE Name	Used Technology	Pros	Cons
1	Hakia	OntoSem (sense repository), QDEX (Query indexing method), Semantic Rank algorithm, Link & Free text	Determines information easily from trusted websites and gives accurate results	cannot index every individual, and needs helo from other search engines
2	Kngine	Uses fuzzy logic algorithm to determine the relations among concepts and the meaning of terms, synonyms, concept search	results appear as image in one form, it used many languages to allow the user to search in parallel. Categorized Results	The silent mode is not allowed and doesn't work in all browsers
3	DuckDuckGo	The results are found by either static data resources or 3rd party APIs. Clustered search, NLP	Does not record privacy of users' information and search for the results by many resources and web crawlers	Lacks video and image search
4	SenseBot	Uses text mining algorithm to identify semantic key concepts of searcher query and parses websites to perform a coherent summary	Summarizes different types of documents	Does not supports some browsers, only works with Firefox as extension, also display the results by Google
5	Swoogle	Indexing semantic, ontologies, OWL, RDF	Uses ontologies and documents for searching and provides metadata of SWDs	indexing of large data and queries are still challenge

3.11. Sensebot

The search process is to analyze web pages and to define the keywords depending on semantic concepts. It provides many of the documents that are related to search term and make summarization of the content of the documents to give the best answer to searcher [25]. The summary brings the best idea of some topic related to the searcher query. This summary is coherent to the searcher. Also, it saves time to provide the best topic related to search term and references to right resources. The engine tries to understand the whole concepts of sentence to give the suitable answer to searcher [18]. The searcher does not need to open many pages to meet his requirements [20].

3.12. Swoogle

Swoogle is an intelligent SSE that searches about the meaning of the words instead of the syntax. In addition, it is considered as a crawler depending on indexing and retrieval systems [24]. The structure of Swoogle is divided into four main components which are: 1) metadata creation, 2) Simple Web Discovery (SWD), 3) data analysis, and 4) the interface [27]. At the backend the SWD creates the database of the SW documents depending on a hybrid approach. It uses the address of URLs, URLs from conventional search, analysis SWDs and generates new URI candidates, this is used to generate URLs to find SWDs on the web. The indexing part is used to index SWDs depending on its metadata. The used techniques in this engine are RDF/XML, N-triple [16]. In addition, some other languages are also used such as OWL, DAML, PDFS and RDF. The analysis part of the engine used to create metadata to describe documents to classify the Semantic Web Ontologies (SWOs) and Semantic Web Databases (SWDBs) [19]. The last part of the system called services which considered as the engine interface that tries to provide data services depending on ontologies at the term level [27].

4. Literature Review

The SW is one of the important subjects that many of authors work within this area. These days everyone focuses on some features and techniques that uses the SSEs. Sahu et al. [12] made a comparison among four kinds of search engines according to their performance, they conclude Google as the best one has features. In addition, they give better results in most cases compared with other type of search engines because it uses the semantic query. Finally, they concluded that Google and Yahoo are developing every day but Bing is developed every month while Ask.com had become every old. The order from top to down: Google, Yahoo and Bing respectively.

Shah et al. [22] compared between different SSEs by using RDF technique in their approach. The RDF is depending on several classification criteria on SSEs. In addition, some technologies that used there are discussed with the evaluation of their performance. The advantage and disadvantage of each search engines are also presented there. The paper included analyzing the search engines' technologies and how a researcher can reduce the flaws for each engine. Finally, discussions for each engine as better, most suitable depending on the purpose, the quality of results in SSEs are presented and how do they need to improve day by day.

Malve and Chawan [18] concludes that the SSEs are better and have many advantages over the KSEs in terms of accuracy of presenting the results. The process of search in SSEs depends on semantic queries. In SW, the users have more assurance to achieve the accuracy of information and getting the answers based on the meaning of words that been searched by users instead the page rank algorithms and keywords. In addition, the main different between KSEs and the SSEs are presented. They also present the clear idea of the techniques used in SW which enables the user to achieve the best information.

Qureshi et al. [19] focus on exploring differences of SW search dimensions. They use the excellent pyramid to test the different dimensions to study about the SSEs. Even now the SSEs are in their developing stage and there are few numbers of resources in this field. In addition, there are many of explore search by querying for various device and difference semantic search engines record and stored formerly. All related materials used in semantic search can be obtained by many authors. Each of the search engines are depending on the pyramid of standard SSE. The authors also compare analysis between different of emerging search engines depending on pyramid which shows the requirements of the search engines.

Jain et al. [17] concluded that, the web 2.0 search engines are different from the SW search engines, because those in the web 2.0 search engines are unable to give the answers directly to user's query. The reasons of this problem are the web 2.0 search engines consist of unstructured information in nature while the web 3.0 (SW) uses RDF format to form the information in more suitable structure which helps semantic search engines such as Falcon, Swoogle, SWSE etc. understand the data and try to give more efficient results to the user. Web 3.0 deals with the data which is structured by RDF or OWL formats only. In addition, the web 2.0 also consists of large library of data linked together in semi-structured such as CSV and XML. The data in web 2.0 can rearrange in OWL format which can be benefited by the SSEs to expand the area of data search. Finally, the authors concluded that SSE technique is what is used in crawling, indexing, ranking and result formation process.

Jagtap et al. [20] briefly surveys many kinds of the SSEs which use different type of methods to search information for the user query. Furthermore, a comparison between the intelligent SSEs and their techniques, and the search engines which depending on high recall of perspectives but low of accuracy are presented. The determination of identify users,

Table 3: Comparison Based on Literature Review.

No	Refs.	Search Engines	Technologies	Purpose of paper	Conclusion
1	[12]	Google, Yahoo, Bing, Ask	OWL, RDF, XML, HTML.	Checks the speed of the question answer with the accuracy of information	Concluded that the Google is the best one that has features and gives a very quick answer
2	[22]	Intelligent Web Service Search, XploreProducts.com platform	RDF, SPARQL DBpedia	Explains the cons and pros, the best approaches selected of the SSEs	Compares among several SSEs according to their approaches and the criteria that used in each one
3	[18]	Wikipedia, Google, DuckDuckGo, Hakia, Swoogle	OWL, RDF, HTML, XML	Compares between KSEs & SSEs	The SSEs are better than KSEs and the results in SW is more accurate in results
4	[19]	Hakia, Lexxe, Cognition, Powerset, Sensebot	QDEX algorithm, NLP, RDF, RDQL	Compares among several SSEs according to various criteria such as Environment, Query Type, Intrinsic Problem	The SSE in the developing stage and they depending on the pyramid standard of SW levels
5	[17]	Falcon, SWSE, Swoogle, Hakia, Google	URIs, RDF, SWDs, OWL, XML, CSV	Explains the main features of several SSEs and the evaluation of each one	The search engines in Web 3.0 are better than those in the Web 2.0, and the SW is quicker to get the answer and the results related to user's request
6	[20]	MSN, Google, Proposed system, Yahoo	OWL, RDF, QDEXing, XSEarch, XCDSearch	Compares among several SSEs according to recall rate and precision rate	Concluded that the SSEs have different tools and techniques to search about the information with high recall rate and low accuracy
7	[21]	Google, DuckDuckGo, Kngine, Hakia	RDF, DAML+OIL, OIL and OWL, XML, SEWISE	To explain four approaches where the SSEs depend on, and the techniques that used of each one.	Concluded that some of SSEs use different approaches to decrease the exclusive of the use and to achieve to better information relatives to the user request
8	[22]	Swoogle, Factbites, DuckDuckGo, Sensebot, Kngine, Lexxe	RDF, OWL, QDEX (Query indexing method), fuzzy search, Semantic Rank algorithm	Presents a comparative study among techniques used in SSE	Searching within the internet is a challenge and needs new approaches to improve utilization of search engines

inaccuracy queries and crawler efficient, and the used tools in the SSEs for search of information on the websites are discussed as well. In addition, the development of SSEs is efficient and uses the technology to answer the complex queries of users. Also, the author makes short overview of the best SSEs which use different approaches in many methods to present the unique search experience for the users. The search on the internet today is a challenge because the most of the complex question unanswered while the SSEs present the suitable answers to user's queries.

Chitre in [21] presents some of the SSEs depending on different approaches using by different methods to decrease the exclusive search experience for the users. In addition, the search process on the internet today is challenging to predicate the efficient answers of the user queries that are suitable for the meaning. The author provides ways for how SSE can do better performance outperforms the limitations of the KSEs.

5. **Discussion**

In this section, a discussion for the reviewed search engines will be presented. As shown from table 1, the main differences between SSE and KSE is that, KSE is based mostly on conventional technologies such as HTML, XML... etc. while SSEs are using Semantic Web technologies such as RDF, OWL…etc. As given in table, all of the systems use RDF, OWL and other semantic web technologies. As it is obvious from Table 2, the mutual advantage of the SSE is that all of them retrieve accurate results to the query, while the common disadvantage is that indexing of large chunks of data is a challenge. KSEs are simple to implement and most of the users know how to use it, while SSEs are more complex systems and need the technology of Semantic Web to be implemented. As shown from table 3, most of the SSEs are not yet common to the public, this is due to a good knowledge is required for lay-users to let them use it which makes it complex for them. The giant search engine companies such as Google and Bing started moderately to include Semantic Web in its search results so that the SW concept becomes more familiar for normal users.

6. **Conclusion**

The current Web offers an easy way to share information online, this makes the size of data on the web become huge gradually. Search engines help users to find information on the web. There are two types of search engines: Keywords Search Engines (KSE) and Semantic Search Engines (SSE). The KSE considers as the base search engines of the web, but they cannot find exact and accurate information to the user queries because they depend on the syntax of the keywords. SSEs solve this problem by looking into the meaning of the keywords and retrieves related results semantically. SSE depend on the technology of the Semantic Web which help them understand the concepts and help machines to

understand and process information. In this paper, an overview of Web technology is given, search engines and their types are presented, then a comprehensive comparison for the two most common search engines KSE and SSE are exposed. Finally, a wide-ranging discussion for each of the reviewed systems with their used technologies, techniques, methods, pros and cons are presented in details.

References

[1] K. Jacksi, S.M. Abass, "Development History of the World Wide Web."

[2] A. Gyrard, M. Serrano, G.A. Atemezing, "Semantic web methodologies, best practices and ontology engineering applied to Internet of Things," IEEE: 412–417, 2015.

[3] K. Jacksi, S.R.M. Zeebaree, N. Dimililer, "LOD Explorer: Presenting the Web of Data," International Journal of Advanced Computer Science and Applications (IJACSA), **9**(1), 2018, doi:10.14569/IJACSA.2018.090107.

[4] K. Jacksi, N. Dimililer, S.R. Zeebaree, "State of the Art Exploration Systems for Linked Data: A Review," International Journal of Advanced Computer Science and Applications (IJACSA), **7**(11), 155–164, 2016, doi:dx.doi.org/10.14569/IJACSA.2016.071120.

[5] A. AL MUQRISHI, A. Sayed, M. KAYED, "CASENG: ARABIC SEMANTIC SEARCH ENGINE.," Journal of Theoretical & Applied Information Technology, **75**(2), 2015.

[6] S.R. Zeebaree, A.-Z. Adel, K. Jacksi, A. Selamat, "Designing an ontology of E-learning system for duhok polytechnic university using protégé OWL tool," Journal of Advanced Research in Dynamic and Control Systems, **11**(05–Special Issue), 24–37, 2019.

[7] K. Jacksi, N. Dimililer, S.R.M. Zeebaree, "A Survey of Exploratory Search Systems Based on LOD Resources," in: , in PROCEEDINGS OF THE 5TH INTERNATIONAL CONFERENCE ON COMPUTING & INFORMATICS, UNIV UTARI MALAYSIA-UUM, COLL ARTS & SCI, INFOR TECHNOL BLDG, SINTOK, KEDAH 06010, MALAYSIA: 501–509, 2015.

[8] A. AL-Zebari, S.R.M. Zeebaree, K. Jacksi, A. Selamat, "ELMS–DPU Ontology Visualization with Protégé VOWL and Web VOWL," Journal of Advanced Research in Dynamic and Control Systems, **Volume 11**(01–Special Issue), 478–485, 2019.

[9] D. Fensel, "Ontology-based knowledge management," Computer, **35**(11), 56–59, 2002.

[10] A. Akerman, J. Tyree, "Using ontology to support development of software architectures," IBM Systems Journal, **45**(4), 813–825, 2006.

[11] J. Rashid, M.W. Nisar, "A Study on Semantic Searching, Semantic Search Engines and Technologies Used for Semantic Search Engines," International Journal of Information Technology and Computer Science (IJITCS), **8**(10), 82–89, 2016.

[12] S.K. Sahu, D. Mahapatra, R. Balabantaray, "Comparative Study of Search Engines In Context Of Features And Semantics," Journal of Theoretical and Applied Information Technology, **88**(2), 210–218, 2016.

[13] Z.N. Rashid, S.R. Zebari, K.H. Sharif, K. Jacksi, "Distributed Cloud Computing and Distributed Parallel Computing: A Review," IEEE: 167–172, 2018.

[14] R. R. Zebari, S. R. M. Zeebaree, K. Jacksi, "Impact Analysis of HTTP and SYN Flood DDoS Attacks on Apache 2 and IIS 10.0 Web Servers," 156–161, 2018, doi:10.1109/ICOASE.2018.8548783.

[15] R.R. Zebari, S.R. Zeebaree, K. Jacksi, H.M. Shukur, "E-Business Requirements For Flexibility And Implementation Enterprise System: A Review."

[16] S.C. Damle, S.S. Kupekar, K. Deulkar, "Comparative Study of Semantic Search Techniques using RDF," International Journal of Computer Applications, **975**, 8887, 2015.

[17] R. Jain, N. Duhan, A. Sharma, "Comparative study on semantic search engines," International Journal of Computer Applications, **975**, 8887, 2015.

[18] A. Malve, P. Chawan, "A Comparative Study of Keyword and Semantic based Search Engine," International Journal of Innovative Research in Science, Engineering and Technology, **4**(11), 11156–11161, 2015.

[19] M.M. Qureshi, B. Asma, H.U. Khan, "Comparative analysis of semantic search engines based on requirement space pyramid," International Journal of Future Computer and Communication, **2**(6), 562, 2013.

[20] D. Jagtap, N. Argade, S. Date, S. Hole, M. Salunke, "Implementation of Intelligent Semantic Web Search Engine."

[21] N. Chitre, "Semantic Web Search Engine," International Journal Of Advance Research In Computer Science And Management Studies, **4**(7), 2016.

[22] V. Shah, A. Shah, K. Deulkar, "Comparative Study Of Semantic Search Engines," International Journal Of Engineering And Computer Science, **4**(11), 14969–14972, 2015.

[23] S.R. Ranganathan, M. Marrikkannan, S. Karthik, "NEED OF FUTURE WEB TECHNOLOGY THE SEMANTIC WEB A BRIEF SURVEY," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), **1**(10), pp-71, 2012.

[24] M. Rahman, "Search engines going beyond keyword search: a survey," Int. J. Comput. Appl, **75**(17), 1–8, 2013.

[25] S.S. Laddha, P.M. Jawandhiya, "Semantic search engine," Indian J. Sci. Technol, **10**(23), 01–06, 2017.

[26] G. Sudeepthi, G. Anuradha, M.S.P. Babu, "A survey on semantic web search engine," International Journal of Computer Science Issues (IJCSI), **9**(2), 241, 2012.

[27] S.K. Sahu, N. Kumar, S. Allawadhi, J. Singh, "Comparative Analysis of Semantic and Syntactic Based Search Engines," Journal of Advanced Computing and Communication Technologies, **4**(2), 22–27, 2016.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 108-114 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Discriminant Analysis of Diminished Attentiveness State Due to Mental Fatigue by Using P300

Kosuke Fujita[1], Kazuya Miyanaga[2], Fumiya Kinoshita[*,2], Hideaki Touyama[2]

[1]*DENSO TECHNO Co., Ltd., Aichi 474-0025 Japan*

[2]*Department of Electrical and Computer Engineering, Faculty of Engineering, Toyama Prefectural University, Toyama 939-0398 Japan*

ARTICLE INFO

Article history:
Received: 25 September, 2019
Accepted: 28 October, 2019
Online: 22 November, 2019

Keywords:
Mental fatigue
Electroencephalogram
P300
Linear discriminant analysis

ABSTRACT

Fatigue is broadly divided into two types depending on the content of a task: physical fatigue and mental fatigue. Mental fatigue is associated with human error. It is thus important to search for indicators that can easily evaluate mental fatigue. The aim of this study is to construct a system that can evaluate mental fatigue in a simple manner. To achieve this, we investigated the influence of the accuracy of linear discriminant analysis in two classes before and after the application of a mental load. In addition, we investigated whether the mental fatigue state can be estimated even when the number of trials for averaging is small, by combining the electroencephalogram power spectral density component and the P300 component. As a result, this combination of the power spectral density component with the P300 component resulted in measured waveforms that exhibited an accuracy of approximately 97%, even when the number of trials for averaging was as small as five trials.

1. Introduction

The technological development of wearable devices in recent years has made it easier to measure biological information. In addition to devices that can measure heart rate or activity level, wearable electroencephalogram (EEG) devices that can measure brain waves have emerged, and their use is becoming more widespread. In the past, expensive equipment was generally used to measure brain waves and it was difficult to continually take these measurements during a person's daily life. However, as wearable EEG devices have become more widespread, dry electrodes that do not require the application of conductive paste have emerged. This development is making it more and more feasible to collect big data in the field of neuroscience. However, the data obtained from wearable devices differ from biological information measured by medical devices. As such, these data need to be applied appropriately. This paper focuses on the use of EEG to evaluate mental fatigue as a method of applying measurement data obtained from wearable devices. This paper is an extension of work originally presented in IEEE International Conference on Systems, Man, and Cybernetics (SMC) 2018 [1].

Fatigue is broadly divided into two types depending on the content of a task: physical fatigue and mental fatigue. While the *in vivo* mechanisms and processes of physical fatigue have been studied, particularly in fields like exercise physiology [2], there is no consensus regarding what mechanisms induce mental fatigue or how it should be measured, and even recent studies have taken a variety of different approaches to researching the subject [3, 4]. In the fields of ergonomics and occupational health, the flicker test is widely used to evaluate mental fatigue [5, 6]. The flicker test evaluates eye fatigue by measuring the frequency or flicker interval threshold, also known as the Critical Fusion Frequency (CFF), as the frequency of a flickering light is changed. However, the CFF is self-reported by subjects and the influence of subjective factors cannot be ruled out. Other than the flicker test, attempts have also been made to evaluate mental fatigue using salivary amylase [7, 8]. A rise in salivary amylase is generally thought to indicate an increase in sympathetic nervous activity. However, sympathetic nervous activity can arise because of negative emotions or mental stimulation. As such, evaluating mental fatigue using autonomic nervous activity indicators such as sympathetic nervous activity or vagal nervous activity often leads to difficulties with interpretation. Meanwhile, EEG waveforms contain both components for evaluating autonomic nervous activity and event-related potential components, which can both be used to evaluate cognitive function. This means that it may be possible to estimate the decreased attentiveness that accompanies mental fatigue by combining these EEG components. This paper focuses on brain waves as an indicator for evaluating mental fatigue.

* Fumiya Kinoshita, Email: f.kinoshita@pu-toyama.ac.jp

In previous studies, there are reports in which α-wave band power increases during mental fatigue [9, 10], there are also reports showing that it decreases as well [11, 12], and thus, there is no consistent trend. P300, an event-related potential, is the positive potential change produced 300 ms after exposure to external stimuli. Its amplitude is significant in that the amplitude is dependent on the amount of processing resources at the perceptual-central level assigned to the event that caused the positive potential change. Prolonged latency and lowered amplitude during mental fatigue have been reported [13, 14]. When calculating P300, the signal-averaging technique must be used in order to produce a clearer potential response to stimuli. Generally, it has been confirmed that about 20 trials are required for averaging to produce consistent results for P300 [15], but it is difficult to conduct 20 trials during the performance of actual tasks. As such, for the simple measurement of mental fatigue using P300, the evoked potential needs to be extracted with averaging over as few trials as possible.

This paper presents work that combined EEG frequency components and P300 components to perform linear discriminant analysis on EEG waveforms. Analysis was performed before and after a task that involved a mental load in order to study the number of trials required for averaging and the resulting discriminant accuracy. Furthermore, in order to confirm that the task had caused sufficient mental fatigue, subjective questionnaire surveys and studies of autonomic nervous activity indicators using electrocardiogram (EKG) analysis were also conducted at the same time.

2. Experimental method

The subjects were 10 young, healthy males (mean ± standard deviation: 21.7 ± 0.9 years) with no history of nervous system disorders. The experiment was thoroughly explained to subjects beforehand and their consent to participate was obtained. The experiment was approved (4/2017) by the ethics committee of Toyama Prefectural University.

With subjects seated, the experiment measured mental fatigue with EEG, electrooculogram (EOG), and EKG analyses as well as a subjective questionnaire survey using the Visual Analogue Scale (VAS). The VAS evaluation axis in this study placed "optimal sensation with no fatigue at all" (0%) at the left end and "lowest level of sensation with so much fatigue that I cannot do anything at all" (100%) at the right end [16]. The distance from the point marked by the subject on the evaluation axis to the 0% point on the left end was measured with a ruler, and the value was expressed as a percentage. Biosignals were recorded with a g.USBamp system (g.tec Medical Engineering GmbH, Austria). The time resolution was set to 1 kHz for the EKG and 512 Hz for the EEG and EOG. Based on the extended 10–20 system, electrodes were positioned at Cz and Pz with the GND at AFz and the reference electrode on the left ear lobe. Additionally, EOG electrodes were positioned above and below the left eye, while EKG measurements were taken with lead III positioned between the left leg and left hand.

The experimental protocol was as follows: after applying a mental load for 15 minutes, all signals to be measured were recorded continuously for 10 minutes. One of these trials constituted one set, and a total of 8 sets were performed (Figure 1). Below, results prior to measurements in the 1st set are labeled "Pre" and measurements in the 8th set are labeled "Post." In this experiment, mental load was created with mental arithmetic problems. In these mental arithmetic problems, a pair of two-digit integers was shown on a display placed 0.8 m in front of the subject's eye (Figure 2). The subject mentally calculated the sum of the first and second digit in each integer (in Figure 2, the results were 12 on the left and 8 on the right), then mentally calculated the units digit of the product of the two integers obtained in the previous step (in Figure 2, the answer was 6). The mental arithmetic problem involved the task of entering the resulting number with a keyboard, and this problem produced mental fatigue in the subject.

For bio signal recording, we conducted an oddball paradigm using visual stimuli and recorded EEG, EOG, and EKG data during that time. When the oddball paradigm in each set was complete, we also conducted a subjective questionnaire survey on fatigue using the VAS. In the oddball paradigms using visual stimuli, an "×" image was used as the infrequent stimulus and an "○" image as the frequent stimulus, with the images set to appear 30 times and 90 times, respectively. Both visual stimuli were displayed for 1 second at a time, and the display order was random. In experiments that use oddball paradigms, P300 detection accuracy is often increased by giving subjects prior instructions to calculate the number of times the infrequent stimulus appears. In this experiment, we instructed subjects to "Quickly press the Enter key when the infrequent stimulus is displayed". To synchronize biosignals with the presentation of stimuli during the oddball paradigm, we synchronized biosignals with trigger signals by installing a photoresistor circuit in the stimulus display and sending external inputs to a biological amplifier.

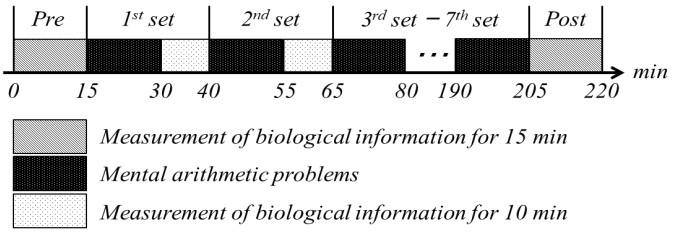


Figure 1: Experimental protocol.

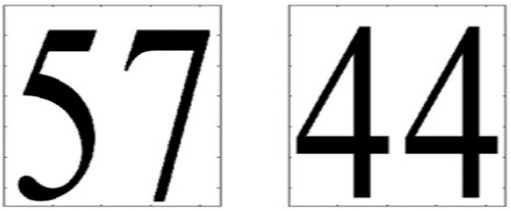


Figure 2: The typical content used in mental arithmetic problems

3. Analytical method

3.1. EKG analytical method

The R-R interval was extracted from the measured EKG time series data and an analysis was conducted using heart rate variability (HRV) [17]. With HRV, analyzing the R-R interval

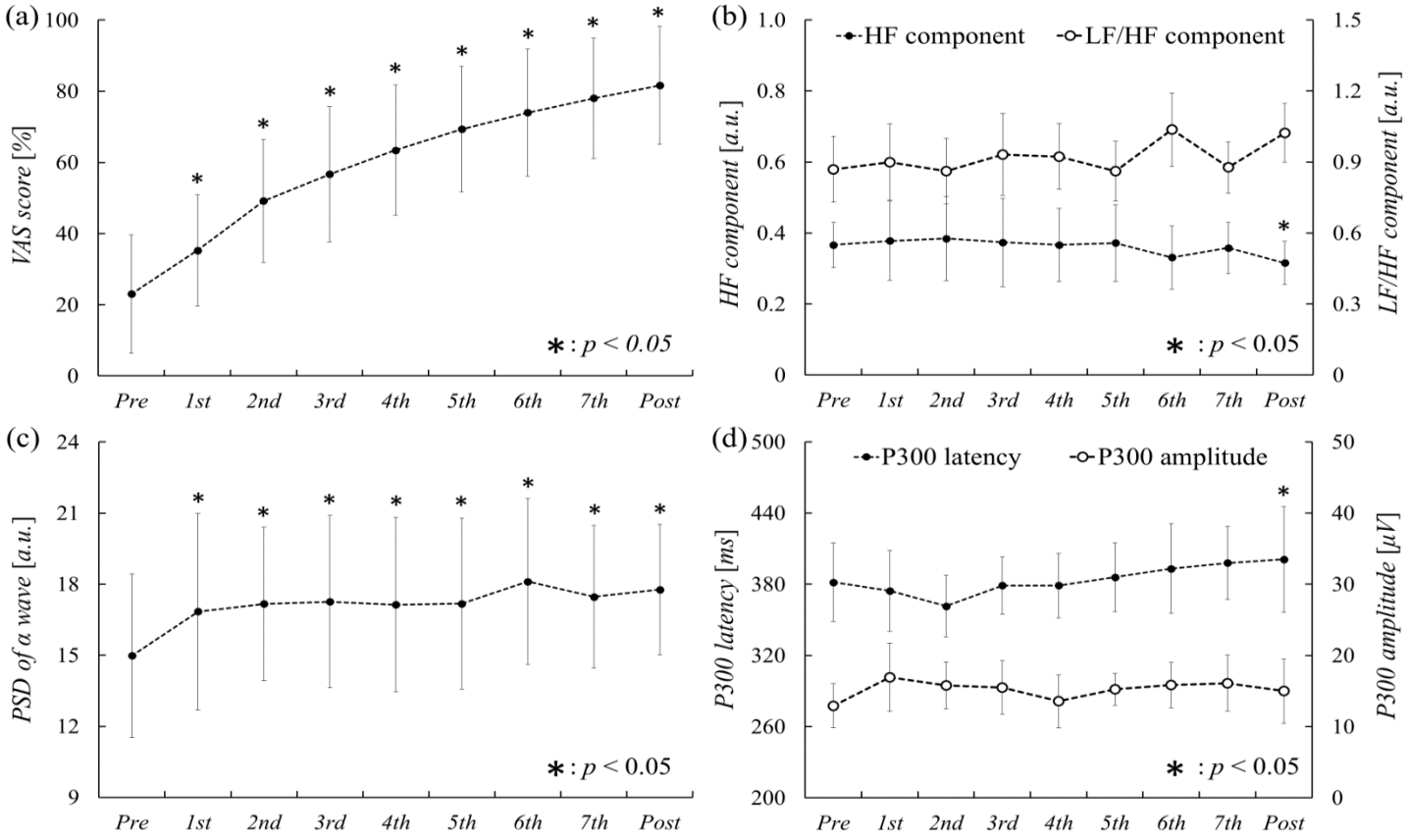


Figure 3: Results of measurement items (Average ± SD)
(a) VAS scores, (b) HF and LF/HF components, (c) PSD of the α wave, (d) The latency and amplitude of the P300 component.

(heart rate) from the time domain/frequency domain makes it possible to quantify sympathetic and parasympathetic nervous system indicators. The extracted R-R interval time series was resampled from a 256-point time series and a fast Fourier transform (FFT) was used to estimate the power spectral density (PSD). Using the PSDs of a low frequency component (LF) with a range of 0.04–0.15 Hz and a high-frequency component (HF) with a range of 0.15–0.4 Hz, this study calculated an LF/HF component and an HF component, which are considered indicators of sympathetic and vagal nervous activity, respectively.

3.2. EEG analytical method

Time series of 5 seconds before and after low-frequency stimulus presentation (a total of 10 seconds) were extracted from the measured EEG time series. After applying a 4–30 Hz bandpass filter to the extracted time series, eye blink artifacts were removed with the FastICA algorithm [18]. Then, data from 2.5 seconds before and after low-frequency stimulus presentation (a total of 5 seconds) were extracted from these time series and put through FFT processing. The PSD of the α bandwidth (8–13 Hz) in each time series and the average values from Cz and Pz were calculated to produce representative values for the sets.

Next, we will demonstrate how P300 was derived. Time series of 5 seconds before and after low-frequency stimulus presentation (a total of 10 seconds) were extracted. After applying a 1–5 Hz bandpass to the extracted time series, eye blink artifacts were removed with the FastICA algorithm. P300 was derived by first correcting the baseline using the average value from 0.25 seconds before and after low-frequency stimulus presentation in the time series and then averaging the waveforms after low-frequency stimulus presentation in each subject. Additionally, this study identified the positive peak that appeared at 280–500 ms as P300 and calculated its latency and amplitude. The average values for P300 latency and amplitude were calculated from both Cz and Pz, and the results were used as the representative values of the sets.

3.3. Linear discriminant analysis

Linear discriminant analysis was performed using the frequency component of the α bandwidth before and after a mental load task and the P300 component of the time series [19]. For the frequency component and P300 component, principal component analysis was used to perform dimensionality reduction, and feature points were extracted [20]. To prevent overfitting, we calculated all levels of accuracy when the number of components in the principal component analysis was changed in the range of 1–15, and the accuracy level in the scenario that reached the local maximum was set as the final discriminant accuracy.

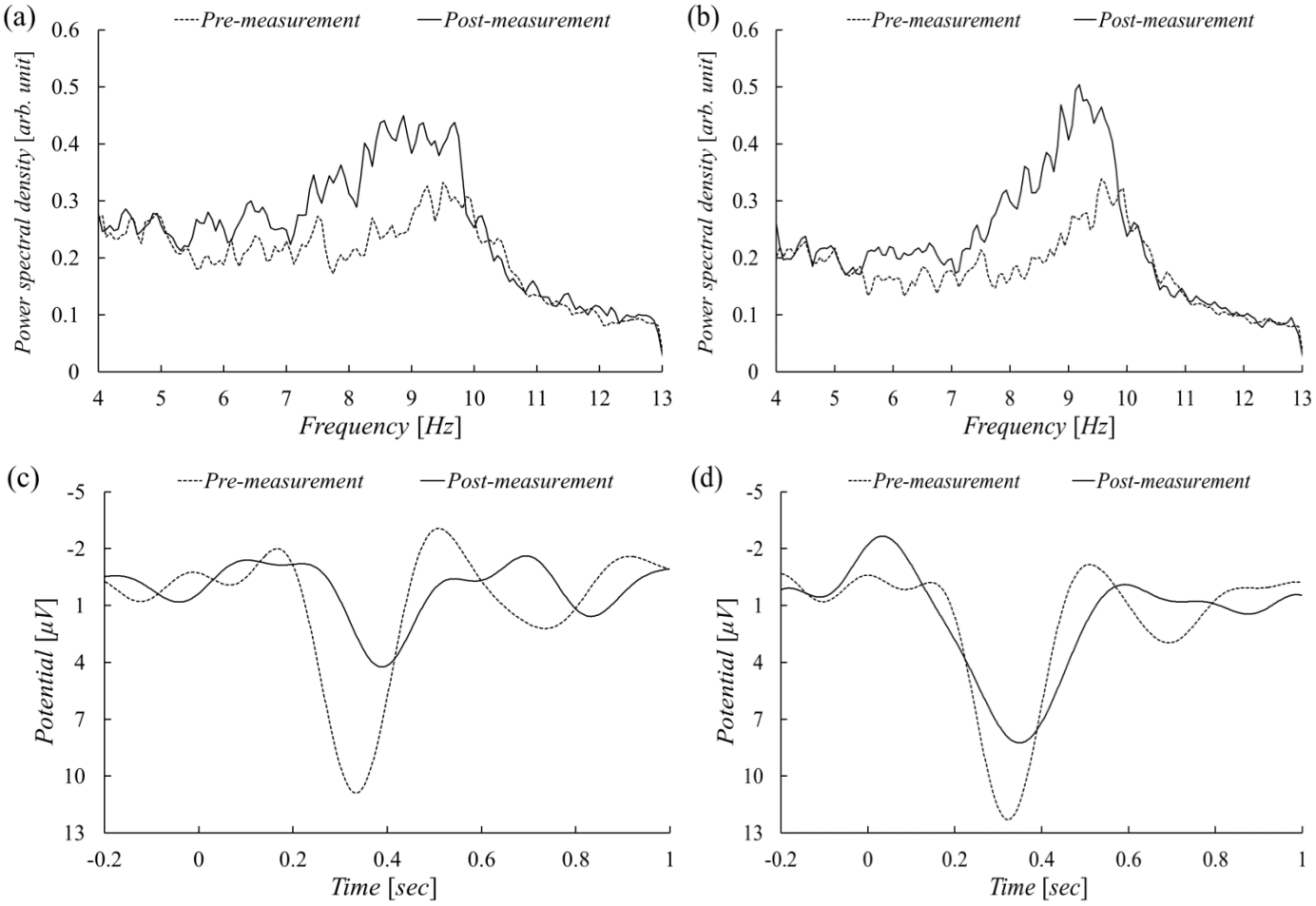


Table 1: Results of Linear Discriminant Analysis (F Value, Average ± SD)

	Power spectrum component		*P300 component*		*Power spectrum component and P300 component*	
Averaging Times	*Pre*	*Post*	*Pre*	*Post*	*Pre*	*Post*
0 *times*	0.74 ± 0.10	0.71 ± 0.09	0.67 ± 0.06	0.68 ± 0.06	0.79 ± 0.08	0.78 ± 0.08
2 *times*	0.79 ± 0.11	0.77 ± 0.12	0.77 ± 0.09	0.77 ± 0.09	0.85 ± 0.11	0.85 ± 0.09
5 *times*	0.83 ± 0.15	0.82 ± 0.15	0.96 ± 0.06	0.97 ± 0.04	0.97 ± 0.04	0.97 ± 0.03

4. Results

Before performing a discriminant, analysis using EEG time series, a subjective questionnaire survey and a study of autonomic

nervous analysis indicators using EKG analysis were also conducted in this experiment to confirm that enough mental fatigue had been produced. Figure 3 shows the changes over time in each measurement item. In the subjective questionnaire survey using the VAS, performing mental arithmetic increased the fatigue score (Figure 3a). As it was a subjective questionnaire survey, there was large variation in VAS scores between the subjects. While the average VAS score was about 20% in the Pre stage, it increased up to about 80% in the Post stage. The Pre results and subsequent measurement set results were compared using the Wilcoxon-signed rank sum test. The results showed that the average VAS score increased significantly in the subsequent measurement sets compared to the score in the Pre stage ($p<0.05$).

For the HF component, which was extracted using heart rate variability and is considered an indicator of vagal nervous activity, the average score was about 0.4 in the Pre stage and about 0.3 in the Post stage (Figure 3b). The Pre results and subsequent measurement set results were compared using the Wilcoxon-signed rank sum test. The results showed that the average HF component score decreased significantly in the Post measurement set compared to the score in the Pre stage ($p < 0.05$). Furthermore, for the LF/HF component, which is considered an indicator of sympathetic nervous activity, the average score was about 0.9 in the Pre stage and about 1.1 in the Post stage (Figure 3b). The Pre results and subsequent measurement set results were compared using the Wilcoxon-signed rank sum test. Results showed no significant difference in LF/HF components between the measurement sets.

For the α bandwidth PDS, the average score was about 15 in the Pre stage and about 18 in the Post stage (Figure 3c). The Pre results and subsequent measurement set results were compared using the Wilcoxon-signed rank sum test. The results showed that the average α bandwidth PDS score increased significantly in the subsequent measurement sets compared to the score in the Pre stage ($p<0.05$).

For P300 latency, the average score was about 380 ms in the Pre stage and about 400 ms in the Post stage (Figure 3d). The Pre results and subsequent measurement set results were compared using the Wilcoxon-signed rank sum test. The results showed that average P300 latency increased significantly in the Post stage compared to the Pre stage ($p < 0.05$). The average P300 amplitude was about 15 μV in both the Pre and Post stages (Figure 3d). The Pre results and subsequent measurement set results were compared using the Wilcoxon-signed rank sum test. Results showed no significant difference in P300 amplitude between any of the measurement sets.

The study presented in this paper performed linear discriminant analysis that combined the frequency component and P300 component to investigate whether mental fatigue can be estimated even when averaging is only performed over a small number of trials. Figure 4 shows a typical example of the frequency component and P300 component before and after a mental load task when averaging was performed over five trials. We observed that the frequency component on the α bandwidth tended to increase after mental fatigue compared to before mental fatigue (Figure 4a). Additionally, we observed that the P300 component latency tended to lengthen after mental fatigue compared to before mental fatigue (Figure 4b).

Linear discriminant analysis was performed using the frequency component of the α bandwidth before (Pre) and after (Post) a mental load task and the P300 component of the time series. Table 1 shows the accuracy results when averaging was performed over 0, 2, and 5 trials. With linear discriminant analysis, combining the frequency component and the P300 component increased discriminant accuracy for all numbers of trials. When averaging was performed for two trials, the discriminant accuracy was about 78% for the frequency component alone and 77% for the P300 component alone, while combining the frequency component and P300 component increased the accuracy to 85%. Additionally, when averaging was performed over five trials, the discriminant accuracy was about 83% for the frequency component alone and 96% for the P300 component alone, while combining the frequency component and P300 component increased the accuracy to 97%.

5. Discussion

The EEG frequency bands recorded when waking are categorized in theta waves (4–7 Hz), alpha waves (8–13 Hz), beta waves (14–30 Hz) [22]. In addition, frequency drops in the α-wave band occur during states of decreased alertness elicited when falling asleep [23]. However, while there are reports in which α-wave band power increases during mental fatigue [9, 10], there are also reports showing that it decreases as well [11, 12], and thus, there is no consistent trend. While evaluating mental fatigue using autonomic nervous activity alone leads to difficulties with interpretation, adding the P300 component makes it possible to evaluate fatigue based on two indicators: cognitive function and autonomic nervous activity. This study performed linear discriminant analysis that combined the EEG frequency component and the P300 component, an event-related potential, to investigate whether mental fatigue can be estimated even when averaging is only performed over a small number of trials. The results showed that, when the P300 component was averaged for five trials, discrimination with 97% accuracy was possible before and after a mental load task. Furthermore, even when averaged over two trials, combining the frequency component and the P300 component produced a discriminant accuracy of 85%. We believe that the technique proposed by this paper makes it possible to estimate mental fatigue even when an event-related potential is averaged over a small number of trials.

In order to confirm that enough mental fatigue had been produced by the mental load when estimating fatigue, subjective questionnaire surveys and studies of autonomic nervous activity indicators using EKG analysis were also conducted in the work presented here. The results showed that subjects complained of fatigue, and subjective fatigue according to the VAS score also increased significantly. Heart rate variability showed a significant decrease in the HF component after a mental load task. There are reports that fluctuation in the HF component may reflect mental load, and the results of this experiment provide more supporting evidence that this is the case. Based on this result, we believe that, in this experiment, subjects were more affected by mental load than physical load. Additionally, this paper presents results that confirmed an increase in the α bandwidth power after a mental

load task, demonstrating the same patterns in changes to basic rhythms during mental fatigue caused by mental arithmetic that Trejo et al. reported [24]. In terms of P300, we confirmed prolonged latency after a mental load task, demonstrating the same trends in changes to P300 during mental fatigue caused by a driving simulation reported by Zhao et al. [25].

In this experiment, keys on the keyboard were only pressed for the low-frequency stimulus in an oddball paradigm where P300 was measured. As such, the possible influence of exercise-related potentials (for example, a decrease in the power of mu-rhythms around 10 Hz) should be investigated. Generally, C3 and C4 are considered the electrode positions most likely to detect exercise-related potentials due to hand motion. The electrode positions used in this experiment were Cz and Pz, and the key pressing was the same throughout all of the trials. Furthermore, we applied a 1–5 Hz bandpass filter during P300 signal processing, and we believe that hand motion had no effect on the results. However, P300 is a signal that tends to be drowned out by spontaneous brain waves. As spontaneous brain waves appear randomly, it has been proven that they will negate each other when averaged, allowing hidden P300 signals to appear. It has been confirmed that discriminant accuracy for P300 increases as the number of trials used for averaging is increased, meaning that the results in this paper are not accidental and can be reliably attributed to the appearance of P300. On the other hand, the linear discriminant analysis in this paper showed 97% accuracy of discriminant using only the P300 component with 5 times averaging process. This is a very high accuracy. However, in this experiment, it was a case of the accuracy where mental fatigue was given sufficiently, so in the future, we will examine the discrimination accuracy according to the degree of mental fatigue.

Lastly, the amplitude of P300 has been shown to depend on subjective probability and stimulus meaning in an additive manner, and the dimension of information transmission has a multiplicative effect on this [26]. As such, teaching data for machine learning needs to be recorded in advance, and it would take an enormous amount of time to reproduce mental fatigue in the manner discussed here. Transfer learning has been proposed as a method for shortening the process of securing the required teaching data [27]. Transfer learning is a technique in which training data from someone else is used for discriminating one's own data. In the future, to improve discriminant accuracy, we will propose a supervised spatial filter that uses training data from others by assuming a similar spatial distribution in the EEG waveforms of different subjects during mental fatigue using transfer learning.

6. Conclusion

Cognitive deficits brought on by mental fatigue include decreased alertness, difficulty solving problems, and lowered situational awareness. There are claims that human error is related to the cognitive deficits associated with mental fatigue, and it is important to search for indicators that can evaluate the decreased cognitive function associated with mental fatigue. This paper presents a study, which performed linear discriminant analysis that combined the EEG frequency component and the P300 component, an event-related potential, to investigate whether mental fatigue can be estimated even when averaging is only performed over a small number of trials. Generally, it has been confirmed that about 20 trials are required to produce consistent results for P300, but the method proposed in this paper was able to achieve a discriminant accuracy of 97% before and after mental fatigue, even when P300 was only averaged over five trials. Moving forward, we will also investigate the detection of mental fatigue in real time during driving and actual tasks by using wearable EEG devices with dry electrodes that do not require the application of conductive paste.

References

[1] K. Fujita, F. Kinoshita, H. Touyama, "Detection of Cognitive Decline Due to Mental Fatigue Using Electroencephalogram" in 2018 IEEE SMC: Passive BCIs, Miyazaki, Japan, 2018.

[2] Japanese Society of Fatigue Science Clinical evaluation guidelines on designated health foods for physical fatigue in healthy people with no pathological fatigue, Journal of Clinical and Experimental Medicine, 228(6), pp.743-746, 2009.

[3] S. Tajima, K. Yamaguchi, H. Kuratsune, Y. Watanabe, "Measuring fatigue with physiological biomarkers" Adv. Anti. Aging. Med., **6**(3), 329–334, 2010.

[4] Y. Tanaka, S. Wakida, "Biomarkers of stress and fatigue" J. Pharmacol. Sci., **137**(4), 185–188, 2011.

[5] K. Hashimoto, "The physiological meaning of flicker values and issues measuring them: flicker test theory and the reality" Jpn. J. Ind. Health, **5**(9), 563–578, 1963.

[6] K. Hashimoto, "Testing for mental fatigue" The Japanese Journal of Ergonomics, **17**(3), 107–113, 1981.

[7] M. Yamaguchi, T. Kanemori, M. Kanemaru, Y. Mizuno, H. Yoshida, "Can salivary amylase be an indicator for estimating stress?" Japanese Journal of Medical Electronics and Biological Engineering, **39**(3), 234–239, 2001.

[8] K. Nomura: The relationship between performance on mental arithmetic tasks and physiological stress responses, Journal of Osaka University of Economics, 68(5), p133-143, 2018.

[9] M. Yamazaki, M. Matsuura, "Adult and elderly electroencephalograms", Japanese Journal of Clinical Neurophysiology, **42**(6), 387–392, 2014.

[10] The Japanese Society of Sleep Research Computer Committee, ed.: "PSG chart for learning: Interpretation method and explanation of sleep polygraph records," The Japanese Society of Sleep Research, 1999.

[11] L. J. Trejo, K. Kubitz, R. Rosipal, R. L. Kochavi, L. D. Montgomery, "EEG-based estimation and classification of mental fatigue," Psychology, **6**(5), 572–589, 2015.

[12] M. Ishibashi, T. Yoshida, "The relationship between brain wave fluctuation and reaction times with the decrease in arousal level," Ergonomics, **36**(5), 229–237, 2000.

[13] N. Okamura, "How the event-related potential P300 is affected by mental fatigue caused by calculation tasks performed for a long period of time" J. Occup. Health, **49**(5), 203–208, 2007.

[14] T.W. Picton, S.A. Hillyard, "Human auditory evoked potentials. II. Effects of attention" Electroen. Clin. Neuro., **36**(2), 191–200, 1974.

[15] J. H. Cohen, J. Polich, "On the number of trials needed for P300" Int. J. Psychophysiol, **25**(3), 249–55, 1997.

[16] M. E. Wewers, N. K. Lowe, "A critical review of visual analogue scales in the measurement of clinical phenomena" Res. Nurs. Health, **13**(4), 227–236, 1990.

[17] M. Malik, A.J. Camm, "Components of heart rate variability: what they really mean and what we really measure" Am. J. Cardiol.,**72**(11), 821–822, 1993

[18] A. Hyvarinen, "Fast and Robust Fied-Point Algorithms for Independent Component Analysis" IEEE T. Neural Networ., **10**(3), 626–634, 1999.

[19] S. Mika, G. Rätsch, J. Weston, B. Schölkopf, K. Müller, "Fisher Discriminant Analysis With Kernels" Neural Networks for Signal Processing IX: Proceedings of the 1999 IEEE Signal Processing Society Workshop, 1999.

[20] K. Fukui, "Past and Present of Subspace Method: Latest Technology Trend: Theoretical Extensions and their Applications" Information Processing Society of Japan, **49**(6), 680–685, 2008.

[21] V. N. Vapnik, Statistical Learning Theory, Wiley-Interscience, 1998.

[22] H. Watanabe, Y. Koike, N. Sakurai, A. Takahashi, H. Iguchi, "Analysis of electroencephalogram fluctuations during mental work: mental arithmetic, inversion and association loading," Japanese Journal of Electroencephalography and Electromyography, **19**, 253–263, 1991.

[23] M. Osaka, "Effects of mental work and alpha wave peak frequency-associated task difficulty," Japanese Journal of Electroencephalography and Electromyography, **11**, 248–254, 1983.

[24] L. J. Trejo, K. Kubitz, R. Rosipal, R. L. Kochavi, L. D. Montgomery, "EEG-Based Estimation and Classification of Mental Fatigue" Psychology, **6**(5), 572–589, 2015.

[25] C. Zhao, M. Zhao, J. Liu, C. Zheng, "Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator" Accident Anal. Prev., **45**, 83–90, 2012.

[26] H. Shinji, W. Yuki, F. Isato, "On the number of averaged electroencephalography epochs for correct detection with a new multiple probe protocol for P300-based concealed information tests" Journal of the Faculty of Human Cultures and Sciences of Fukuyama University, **14**, 99–106, 2014.

[27] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition" in 2015 ICLR: Computer Vision and Pattern Recognition, 2015.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 111-115 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Model of Automatic Parking House

Radek Guráš*, Miroslav Mahdal

Department of Control Systems and Instrumentation, VSB – Technical University of Ostrava, Ostrava, Czech Republic

ARTICLE INFO

Article history:
Received: 13 December, 2019
Accepted: 05 February, 2020
Online: 09 March, 2020

Keywords:
Automatic Parking House
Vehicle Parking
Storage Mechanism

ABSTRACT

This article deals with the issue of parking houses, storage mechanisms and how to control them. A laboratory model of a multi-storey underground automatic parking house was designed and created for this purpose, which is used to test a new innovative way of storing cars and specialized approaches to driving actuators. The model is based on an aluminum profile frame to which a newly made linear two-stage dual-sided telescopic mechanism and a two-axis positioning mechanism is attached. The emphasis is placed on the smooth movement of all axes which limits shocks caused by mechanism speed change while moving to a defined position.

1. Introduction

This paper is an extension of work originally presented in International Carpathian Control Conference 2019 [1] which deals with the design of whole automated parking house and new solution of car storing.

Automatic parking systems offer modern, sophisticated and intelligent car parking solutions in places where there is a lack of parking space for many vehicles. In such places as city centers, airports, hotels, housing estates, etc. it is possible to use parking houses, which will increase the parking area by the number of floors, whether in the above-ground or underground variant. This solution can be further improved by using an automated system with many advantages over the classic solution.

To demonstrate the functioning of a possible solution of such an automatic parking system, a model of an automatic parking house of the underground structure was created on the principle of a two-axis motion device with a dual-sided sliding telescopic storage mechanism. This solution brings a new perspective on the design of the storage mechanism and develops the widely used principles of automatic car parking, e.g. in [2-4].

The control of this model is provided by a distributed control system consisting of Arduino units in two levels. Arduino Nano microcontroller boards form the zero-order level together with the power modules for control of step and dc motors and represent the analogy of real intelligent systems such as intelligent frequency inverters. The Arduino Mega forms the top level, which represents the master control board of the model. Both levels communicate with each other via UART. The higher level is also ready for connection to the visualization software.

2. Description of Construction

For the purposes of this project, a model of a parking house was designed as physical support for the following solution of the control system and its software. In terms of the optimal ratio between price, size and assembly, the dimensions were chosen so that the device was able to work with 1:24 scale car models.

Some components, particularly the propulsion system, do not entirely match the overall proportions of the parking house model, but they have been used for the fact that they are standard, standardized components whose price is many times lower than the price of atypical prototypes made bespoke.

2.1. Design of the storage mechanism

The basic element of each automated parking house is the mechanism that transports cars from the entry area to the individual storage positions. The most commonly used principle is the transfer of a car on a pallet that the individual carriage systems of the garage convey between themselves. However, in this work, the principle used for transferring the car from the static to the moving part of the system is based on mutually passing support bars (see Figure 1), which support the wheels of the car and during their mutual pass the wheel is transferred from the bars with larger pitch to ones with smaller pitch and vice versa. Based on this principle, a storage mechanism has been developed that functions as a

*Radek Guráš, Department of Control Systems and Instrumentation, VSB – Technical University of Ostrava, Ostrava, Czech Republic, radek.guras@vsb.cz

telescopic sliding dual-sided system with a support element mounted on the second stage of the telescopic device. In this way, the mechanism is able to eject the support element by more than one length of the retracted state, thus enabling the vehicle to be moved to the storage position without the support element being physically disconnected from the storage mechanism.

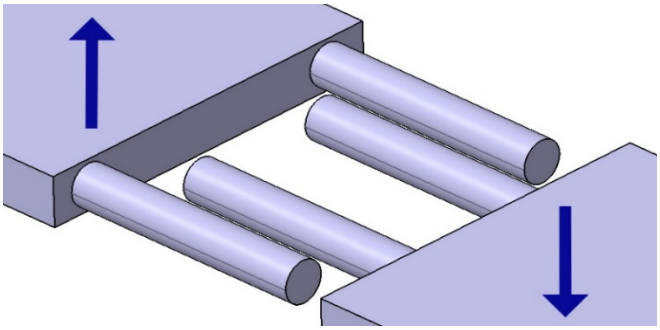


Figure 1: Mutually passing support bars

The storage mechanism (see Figure 2) consists of three basic parts, designed for a linear extension, symmetrically on both sides. The parts were made according to their design from ABS plastic by the 3D printing method based on the requirements for the range of motion, stiffness, and specific dimensions. To measure the deflection of the telescopic structure, a simple strength analysis with a load of 0.5 kg was performed, resulting in maximum stress (according to the HMH hypothesis) of 1.731 MPa and the maximum deflection of the first stage endpoint of 0.66 mm at the maximum assumed load. It has to be emphasized that during the analysis the assembly was considered as a solid body and the properties of 3D printed plastic undoubtedly differed from other manufacturing technologies.

The first telescopic stage drive is realized by a Microcon stepper motor of type SX17-1005LQCEF. This motor has a standardized NEMA 17 flange size, static torque of 0.52 Nm at a rated current of 1 A. The motor movement is transmitted over the pulley by a toothed belt of GT2 type mounted on the opposite end of the first stage slider from the underside.

For the second stage drive, the RC nano-servomechanism was used. It was modified so that the feedback was removed, and the motor is used as a conventional high-speed DC motor with a high ratio gear train due to its small mounting dimensions. The motor is mounted on a second stage moving part and its rotary motion is converted to linear by a pinion and a toothed rack running through the middle of the second stage.

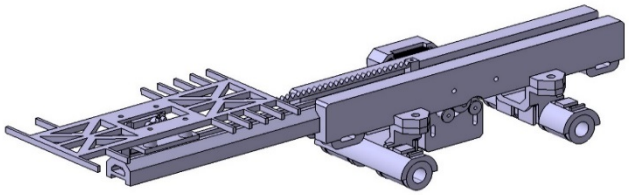


Figure 2: The storage mechanism

2.2. Design of the two-axis motion system

The main system enabling the storage mechanism to move in the handling space is a two-axis system (see Figure 3), capable of moving the car in the vertical axis between the floors and horizontally between the storage boxes within the floor. For this model, the mechanism was designed combining two basic principles: the industrial storage crane and the 3D printer's motion system. Microcon SX17-1005LQCEF and SX17-1003LQCEF stepper motors were chosen to drive both axes. Due to many atypical parts, the method of 3D print was selected to create them. A toothed belt of GT2 type was selected for the transmission of the movement on a horizontal axis. On the vertical axis, due to the requirement for safe stopping of motion in the event of power failure, the trapezoidal screw was used - even at the cost of slower motion.



Figure 3: The motion system

2.3. Design of a frame

As a main load-bearing structure of the entire parking house, a design of a frame made of aluminum modular profiles with dimensions 30 x 30 mm was created (see Figure 4). This solution has been chosen especially with regard to the easy attachment of all parts into the profile grooves.

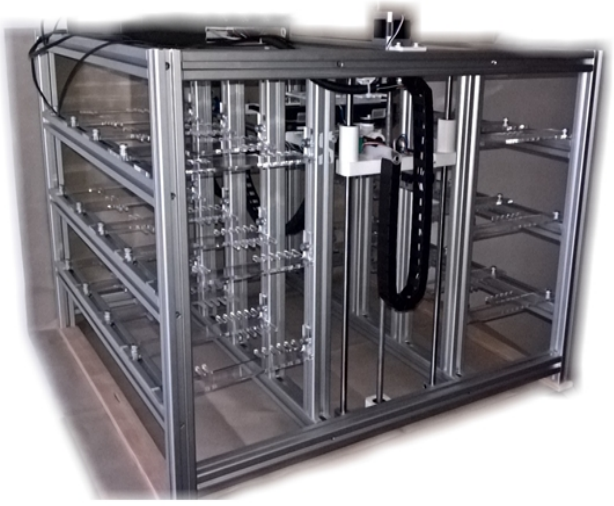


Figure 4: The parking house model frame

3. Realization of the Model

The construction of the parking house model for automatic parking of cars was realized on the basis of a previous design.

3.1. Motion system

The axes are formed by hardened; polished guide rods attached to the base frame. Linear ball bearings embedded in 3D printed carrier parts allow easy motion along the guide rods. Torque transmission of the motors is realized via a flexible coupling, a trapezoidal motion screw and a trapezoidal nut on the horizontal axis and over the pulley and a toothed belt as a drive of the storage mechanism.

The trapezoidal screw of the specific parameters meets the self-locking condition, so even in the event of a power outage or other failure, it will ensure a safe stop of the vertical movement of the mechanism. For verification, the conditions of self-locking in equations (1) and (2) are used.

$$\varphi' = \tan^{-1}\left(\frac{f_z}{\cos\frac{\alpha}{2}}\right) = 8.82^\circ \tag{1}$$

$$\psi \leq \varphi' \tag{2}$$

where ψ is the pitch angle (here from the manufacturer 4°07′), φ' reduced friction angle of the thread, f_z for steel – steel 0.15 and the angle of the thread profile α for a trapezoidal thread is 30°.

3.2. Storage mechanism

The most important and the most complex part of the whole model - the storage system (see Figure 5) was made using the 3D printing method as an assembly consisting of three main carrying parts and many other supporting elements.

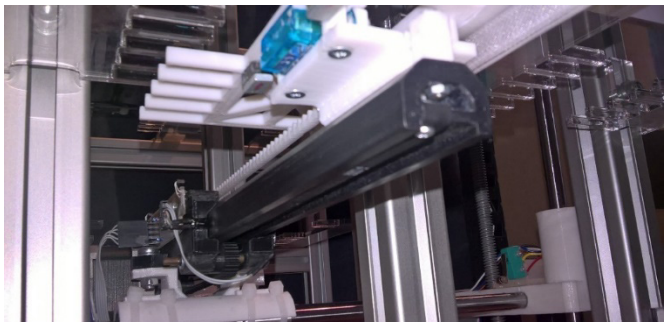


Figure 5: Extruded telescopic mechanism

The main carrier, printed on a 3D printer, has an inner groove for the middle carrying part - the first stage of the telescopic mechanism. Also, there are two smooth pulleys mounted on the underside, through which the toothed belt passes to drive the first telescopic stage. Through the holes on the side of the main carrier, it is possible to mount the stepper motor to drive the first stage.

The second carrier - the first telescopic stage - has such a profile that passes as tightly as possible through the groove in the base part. Also, it has another shaped groove for guiding the moving carriage - the second telescopic stage.

A toothed pinion, made by 3D print, is attached to the output shaft of the second stage motor, transferring motion to the toothed rack, which is solidly connected to the first stage of the telescopic mechanism and passes through the middle of the second stage.

3.3. Static parking spaces

Static parking spaces were made of plexiglass parts cut by a laser (see Figure 6). Ladder shelves have been created, which always in pairs form a parking space for one car. One of their ends was heat-bent at a right angle so that it could be then attached to the main frame of the model.

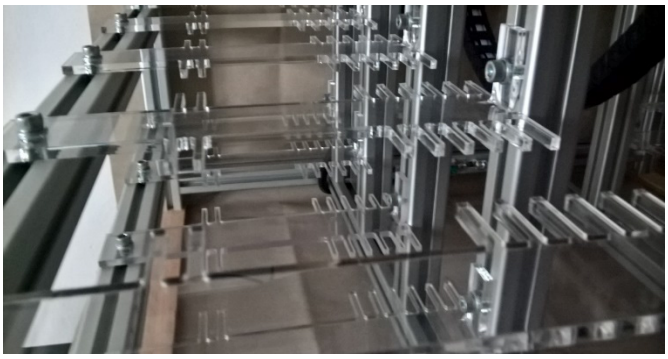


Figure 6: Parking spaces

3.4. Sensors

Several types of sensors were used to build the model. The main purpose of these sensors is to define end positions, homing positions, pullout position control and parking space occupancy detection.

For these purposes, sensors of the following types were used:

- Mechanical end stops (homing, handling space definition)
- Hall sensors (determining the position of the second stage of the telescopic mechanism)
- Optical IR gates (homing of the first stage of the telescopic mechanism, incremental sensors)
- Reflective IR sensors (parking space occupancy detection)

For easier processing of signals from occupancy detection sensors, the PCB was used, with sensitivity tuning and conversion to serial signal using a shift register of "parallel in – serial out" type (see Figure 7).

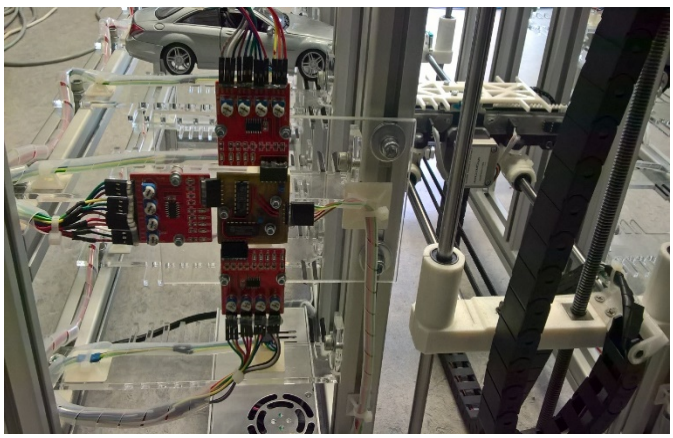


Figure 7: Signal conversion boards

All cabling has been run in power chains in places of motion to prevent cables from entering the moving parts of the system.

4. Model Control

A multilevel system (see Figure 8) has been designed to control the model, consisting of a master controller, build on the Arduino Mega 2560 microcontroller board, which provides data acquisition, communication with the user and with superior control level, and mainly management of three subsystems controlling the movement of each axis of the model. These subsystems are built on the Arduino Nano microcontroller boards, which in combination with the Pololu DRV8825 power stepper motor driver and with the L298N H-bridge for driving DC motor, is an analog of a real intelligent control system, such as an intelligent frequency converter.

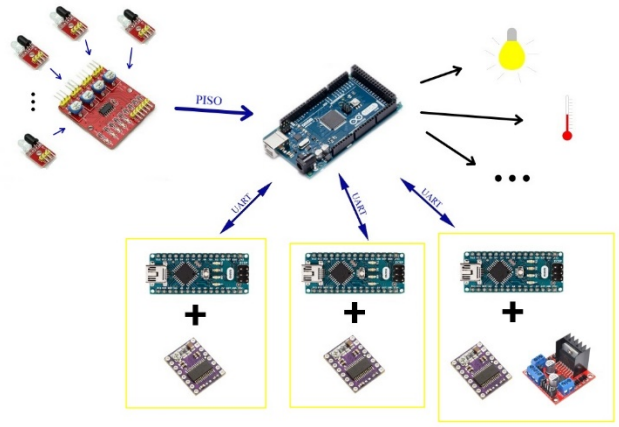


Figure 8: Control system diagram

The UART hardware communication device was used for serial communication between the mainboard and the subsystems.

The DC motor is controlled by a bridge driver by switching the respective inputs which control the power components of the driver and the speed of motion is given by the input PWM modulated enable signal.

4.1. Stepper motor control

A separate chapter in the control algorithm is controlling stepper motors using the Pololu DRV8825 stepper motor driver. This driver switches the individual phases of the motor, eventually provides micro-stepping, depending on the DIR and STEP input signals, where one determines the direction of motion and the other responds to the starting edge of the pulse by performing one step of the motor.

Control pulses for the driver are generated by the microcontroller of the respective axis using the timer interrupt handler. This, depending on the desired time gap between pulses, generates an interrupt flag, triggering the service subroutine, switching the respective outputs and calculates the requested pause time. The longer the pause is, the slower the motor moves.

To ensure the smoothness of the engine's acceleration and deceleration, and to prevent step loss due to a rapid transition from idle to maximum speed, an algorithm was developed to accelerate and decelerate the engine according to the ramp function. Some of these principles can be also found in the online articles (D. Austin, *Generate stepper-motor speed profiles in real time*, Atmel, *Linear speed control of stepper motor*) or in [5].

The time between control pulses can be calculated using the following formula (3)

$$c_n = c_0\left(\sqrt{n+1}-\sqrt{n}\right) \tag{3}$$

where c_0 is the initial pulse pause length, c_n is the length of nth pulse pause and n is the calculation step. This way the ramp function is calculated until the motor speed reaches the maximum value.

Since the calculation of the two square roots is very demanding for the microcontroller processor and does not meet the requirement for the shortest possible interrupt service subroutine time. It is useful to replace the previous formula with Taylor series (4) with a certain, not so significant inaccuracy, and calculate the pause between steps from the previous value according to

$$c_n = c_{n-1} - \frac{2c_{n-1}}{4n+1} \tag{4}$$

This measure reduces the calculation to only division of two numbers that is considerably faster than the square root calculation. In order to further accelerate the calculation, measures have been introduced replacing the more demanding division of decimal numbers by dividing integers into two phases. Firstly, the integer is divided normally and then the modulo division is used to obtain the remainder of the division, which in the next step is added to the divider to reduce the inaccuracy of the calculation [6].

4.2. Homing

Movement of motional parts to the default, home position called "homing" is accomplished by a series of algorithms, each controlled by the respective combination of the microcontroller and the power driver.

The horizontal and vertical axes travel to its home limit switch at a reduced speed, and after activation, it will leave enough to switch off the switch again. The algorithm also considers the situation when it is necessary to get into the home position after a non-standard situation, when one of the limit switches is active at the beginning of the process, especially when the home switch is active. In this case, the switch will firstly move from the switch and then the homing will start.

The first stage of the storage mechanism performs homing by means of two optical gates located at the end of the base part of the mechanism. If the two gates are not interrupted, the slider moves in small increments in the direction of the uninterrupted sensor. At the moment of its deactivation, a predefined deviation from the default value is corrected.

The second stage firstly "moves back" to the nearest Hall sensor in the given direction and then moves forward to its home

sensor. If the stage moves back even to the outer sensor, it is assumed that the starting position is already verified with enough space behind it.

5. Conclusion

During a realization of this project, a model of an underground parking house for parking cars, driven by stepper and DC motors, and designed with suitable position sensors, end position sensors, and sensors of occupied parking space, was designed and built. Many 3D prints have been used to produce many atypical parts. The individual parking positions are made of plexiglass ladder shelves. Vehicle storage is performed by a dual-sided sliding telescopic storage mechanism specially designed for this model, operating on the principle of passing support bars.

Next, a multi-level control system based on Arduino microcontroller boards was designed so that one of them serves as a master controller of the model and others formed intelligent drivers for the individual motion axes of the model. All of these individual boards have been programmed and interconnected by the UART communication device.

This project brings a new perspective on car manipulation and storing and allows further development of algorithms for optimal automatic parking.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by the European Regional Development Fund in the Research Centre of Advanced Mechatronic Systems project, CZ.02.1.01/0.0/0.0/16_019/0000867 within the Operational Programme Research, Development and Education and the project SP2020/57 Research and Development of Advanced Methods in the Area of Machines and Process Control supported by the Ministry of Education, Youth and Sports.

References

[1] R. Guráš, M. Mahdal, "Parking house action subsystem control," in Proceedings of 2019 20th International Carpathian Control Conference (ICCC), Krakow-Wieliczka, Poland, 2019, doi: 10.1109/CarpathianCC.2019.8765989

[2] U. Maheswari, D. Aswini, G. Murtugudde, B. Karthiga, "Automatic car parking system" Int. J. Pure Appl. Math. (Online), Vol. 120, No. 6, 2018. ISSN: 1314-3395

[3] I. L. J. Hamelink, The Mechanical Parking Guide, CreateSpace Independent Publishing Platform, 2011, ISBN 1-466-43786-3

[4] A. Mathijssen, A. J. Pretorius, "Verified Design of an Automated Parking Garage" in Formal Methods: Applications and Technology: 11th International Workshop, FMICS 2006 and 5th International Workshop PDMC 2006, Bonn, Germany, 2006, doi: 10.1007/978-3-540-70952-7_11

[5] P. Novák, Mobilní roboty: pohony, senzory, řízení, Praha: BEN - technická literatura, 2005, ISBN 80-7300-141-1

[6] J. Czebe, J. Škuta, "Usage of single-chip computers for control MIMO systems," in Proceedings of 2017 18th International Carpathian Control Conference (ICCC), Sinaia, Romania, 2017, doi: 10.1109/CarpathianCC.2017.7970428

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 112-118 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Internalising Negative Self-Image Externalities: The First Objective for City Marketing as a Municipal Management Tool[1]

Manuel Escourido Calvo*, Antonio Javier Prado Domínguez, Valentín Alejandro Martínez Fernández

Faculty of Economics and Business, A Coruña University, 15008, Spain

ARTICLE INFO

Article history:
Received: 26 November, 2019
Accepted: 04 January, 2020
Online: 15 January, 2020

Keywords:
city marketing
externality
self-image
sense/pride of belonging
perception

ABSTRACT

In times of crisis or traumatic transformation processes, one of the most frequent negative externalities is that the self-image of a city deteriorates among its stakeholders, which affects their economic and social expectations and decisions. The internalisation of this externality must be a key objective for the local public manager so that any initiative can be carried out with collective support and collaboration. When applied to this end, city marketing and the instruments that this technology makes available to city managers provide optimal results.

1. Introduction

From a historical perspective, cities are understood as the domains of human interaction in which civilisation develops; for this reason, they play a role in supporting economic activity [1]. In fact, an economic analysis of civilisation must take into account its concreteness: that is to say, the city itself. This is the reason why many disciplines approached the subject through the study of the city and its management, knowing full well "the city is, manifestly, something very complex" [2].

The importance of "good" city management is manifest, above all, in times of economic crisis. The crisis of 1973 led to the adoption of a new conceptual framework for city management during a time of restructuration, with a need for social policies aimed at job creation, urban renewal policies and economic policies focused on externalities. Years later, the 1990's industrial depression meant that the city had to make city management decisions based on new technologies, innovation and the negative impact externalities had on urban ecosystems.

Urban managers began to acknowledge the importance of remaining competitive in the city global market during the 2008 crisis under the threat of population decline, unemployment or the loss of investment in favour of more competitive cities.

The city is the crossroad where interests and actions meet and where economic, social, urban and environmental conflicts are expressed with great intensity. When this scenario depends on the fate of industry or industrial groups and resources and/or local produce is in decline- in terms of job creation and invigoration or economic engine- the challenge is even greater.

This has happened in many cities that went through and/or have ongoing industrial restructuring or termination/closure procedures related to the natural resources on which they were dependent. In turn their downfall triggers a demographic threat, with the loss of population in the age bracket that has the biggest impact on birth rates. Decline can also be expressed in economic terms with the loss of future, high quality taxable base income, added value workers, entrepreneurs and externalities, as well as in social terms, with an overall loss of talent.

As explained in research by Escourido, M. [3], which is part of his PhD thesis, this environmental change and foundational transformation of the city resulted in an enormous amount of public funding being spent on discretionary actions in an effort to solve a multifaceted problem with only one solution. The problem was coupled with the absence of a systematic and competitive diagnosis of the city and a much needed strategic thinking process focused on both supply and demand. These actions were detached from any planning and had no impact on the city, although political marketing slightly benefited by them.

*Manuel Escourido Calvo, Email: manuel.escourido@udc.es
[1]This paper is an extension of work originally presented in the 2019 14th Iberian Conference on Information Systems and Technologies (CISTI)

The process of change facing the city this century happen in such a quick and constant pace that, in many cases, the community is overwhelmed and unable to react. If there are ongoing, dramatic situations in the city, the challenge is greater. Nevertheless, cities in decline and lacking in expectations also have citizens asking for solutions to problems related to employment, education, safety and social services. Their demands reach urban managers. It is these public servants who must allocate suitable services to meet their needs and improve the ranking of their city. Being unsuccessful in meeting these goals means that locals can "vote with their feet" [4].

We must not forget that cities are made up of the people who define them, adding dimensions and guiding them to their future. This is why problematic areas are not only those with business or industry in decline, but also those where people think the future will be the same as today [5].

We are then looking at city management with certain features. Its social and economic development process originated from public sector activity, such as in a state energy policy and/or industrialization. Moreover, the economic industrial monoculture model drifted because the market and the comparative competitiveness of factors and sectors went into decline. Even if this trend had been caused by market failure, it suggest that the public sector had erred in its passivity [6].

On the other hand, interventionist measures are overlooked or left incomplete when the city is faced with a socioeconomic change of such magnitude that it leads to a series of negative consequences. Among these outcomes are a highly negative externality and an absence of expectations. Another is the negative self-image of the city, which mirrors its dwindling attractiveness and its social and economic decay. These problems can be offset by changing how the city is perceived with marketing based on urban management tools funded with public-private collaboration formulae.

2. Coase and the externality internalization (solution)

State intervention through regulations, bans and taxation is the usual response to the externalities problem [7], always under the wrong However, in doing so, the public sector determines legal responsibilities instead of clarifying the economic nature of the problem. Thus, the system's efficiency may be diminished.

This happens because the State and courts argue over who has the right to take action rather than decide what should be done. These initial legal rights boundaries can always be modified through market transactions, redefining themselves in the process if this leads to increased production.

At this point, one can follow the thinking of Coase. He refers to, but questions, the assumption that there are pre-existing property rights, assigned and defined in a clear manner, with inexistent or low transaction costs. Enshrined in law, these property rights are in conflict with a viable and efficient economic solution. Thus they would be reallocated towards those who value them the most, even if the legal system goes against them.

With zero transaction costs, economic agents would elaborate the contractual ruling necessary to maximise production value, whatever the nature of these circumstances may be: legal, with liabilities or through wealth distribution. At the same time, Coase [8] states that, when real costs are generated and must be accounted for, many of these agreements and deals may be annulled because their cost would be higher than the income generated.

Taking into account transaction costs, Coase points out that both parties may feel their incentives are lessened when the information needed to carry out deals and agreements have been revealed. Public intervention would be needed to define which incentives are missing or which contractual deals can be done. Even in the case of exceedingly high transaction costs, state intervention can protect specific activities that are causing harmful externalities.

In this context, a point may have been overlooked: communal private property- mainly business- is defined as a system aimed at minimizing transaction costs, even more so in complex social environments.

In conclusion, this is a general balance model for a market economy, whose theoretical underpinnings are closer to being defined. They aim at maximizing efficiency, without taking into account the reason or use of property rights. The State still has the right to act when transactions have high cost or pose difficulties. Otherwise, these agreements will be untenable. Another advantage of the Coase Theorem is that the aforementioned situation is suitable for the externalities of production or consumption. It also comes into play in cases of inaction, when a forest has not been maintained properly or if a government fails to foresee the socioeconomic consequences of a declining industrial monoculture.

Nevertheless, Coase warns that the proposition is weakened by the sense of instability in the balance; the economy is defined as a process, not a state.

With these arguments in mind and heeding Cooter and Ulen [9], one may reach the conclusion that the State's main function is be to minimise the cost of every economic transaction, wherever it may be. In most cases, property rights would be clearly defined. This principle, known as Coase's Normative Theorem, means that law is structured "in such a way that nothing prevents private agreements from being reached". Apart from facilitating and promoting deals, it should minimise losses caused by the lack of deals [10].

3. Including the particularities of subjectivism and perception in internalising (solving) externalities.

Some of the objections to Coase's Theorem are related to the concept of human preferences. These objections, like human preferences, can be subjective, dynamic, shifting and therefore not measurable or comparable. Every market agent would try to increase their "psychological profits" in search of solutions that make them possible. This is mostly a problem of information and perception, as well as a strategic game and aversion to risk. Even taking zero transaction costs into account and properly defining private property rights, the potential possibilities of each deal would have to be taken into account to solve an externality.

As stated by Hayek [11], social science measures human perception around the world in terms of people's beliefs. A true or false perspective cannot be observed directly; the speech and

actions of an individual, however, can be studied. Human actions are not an objective fact: they are what the person performing them believes those action do, and so they form part of a personal, individual perception. Therefore, the perception of goals and means is introduced as an element to consider when solving externalities.

The same applies to cost and risk perception because the elements of individual choice depend on new ideas and are taken into consideration by the person who decides them and his or her preferences [12]. It is worth noting that, in a current behavioral analysis of Coase's Theorem, one of the results, among others, suggest that, when an "unfair offer" is presented, the desire to break negotiations is so strong that the Theorem does not apply, even when it is against the self-interests of the individual rejecting the offer [13].

It is therefore possible to ascertain that social cost and social products are not objective concepts; different individuals consider benefits and prejudices based on their own perceptions, presumptions and analysis. This means efficiency in the choices made cannot be assessed in an objective way.

At this point, the market has tools- including profit and competition as motives- which provide the incentives needed to define and then meet the needs and desires of consumers. The successful provider earns profits, supplanting the others who fail at this task. These tools do not guarantee that the correct choices are always made, because they cannot be known. However, in the long run, it is expected that these same tools optimise the social net benefit [14].

Public sector management must step in at this point so that it can identify, prioritise, and meet the needs and demands of clients or those managed. To this end, marketing as an urban management tool should be adopted. With the paradigm shift ushered in by New Public Management, the administration focuses on the "client-citizen". The end result is custom-made solutions for citizens, not only in the domain of goods and services but also in the promotion and creation of incentives, particular conducts and desirable perceptions, such as those contributing to the internalisation of specific externalities.

4. Externalities of negative self-image. As Pontes de García Rodríguez (Spain) as a case in point

To favour industrialisation and empower energy production, state economic intervention during the 1970's in Spain favoured lignite mining in As Pontes de García Rodríguez (A Coruña, Galicia, NW Spain). This was done in an open pit mining operation. The mineral was then used as fuel in a power plant that was built by the public sector in the same city. A public company was responsible for the project: Endesa (Empresa Nacional de Electricidad SA). In what had once been a rural city, a period of demographic and economic expansion began. It had a striking impact on the income and wealth of the area.

On the one hand, expropriations were carried out to establish the mining operation, which provided unimaginable wealth to mostly rural land owners with forests, crops and livestock. In many cases, this change also meant access to what resembled public employee posts in the mines, power plant or the services attached to them. Moreover, salaries were well above the average national wage, with complements and benefits: the state paid for accommodation, schooling for employees' children and workers' personal electricity bills. In the same fashion, the income generated by this company in the public treasury meant the council budget continued to grow. As a result, an infrastructure was created to support services whose quality was higher compared to that in other cities of similar size and population.

However, all the mining operations and power plant activity overshadowed local businesses due to a difference in employment opportunities and salaries offered by Endesa. A great amount of resources was allocated towards this company. The primary sector, in contrast, was practically reduced to subsistence farming of crops and livestock in those places where the land was not expropriated by the state for mining. On the other hand, the construction sector developed small companies driven by the high demand for housing for the newly arrived workers for the mines and power plant.

Nevertheless, this development was limited by Endesa itself as it was a public company back then. Endesa built over a thousand homes for employees and their families in exchange for a symbolic rent. At the same time, an auxiliary and service industry for repairs, maintenance and supplies emerged, linked to and dependant on mining and energy production. This sector had a presence in commerce, thanks to many catering services.

This situation of extreme economic dynamism in the municipality and surrounding areas and the positive externality resulting from public sector intervention made a complete turnaround after two events. Firstly, the mining operation was closed in December 2007 in accordance with European Union directives. The second event is the worldwide crisis, as experienced in As Pontes de García Rodríguez, especially from 2012 onwards. The shift in focus of environmental legislation in the European Union conveys a negative externality leading to the direct closure of the mining operation in this municipality. Even more, it was known that lignite exploitation would be curtailed and public intervention was limited or non-existent so little effort was made to minimise the risks arising from a social and economic industrial monoculture with a limited lifespan.

Since then the city has felt the effects of socioeconomic decline: unemployment, retirement and early retirement from the mining operation, a lack of opportunities for the younger population brackets with higher education and the failure of re-industrialisation initiatives despite having the advantages of heavy public funding. Moreover, in the absence of planning and management, the different public administrations have been unable to turn the tide in a place that was "killed by its own success".

In turn, the city faces a demographic threat with a loss of population in the age bracket that has the biggest impact on birth rates; an economic threat with loss of future high quality taxable base income, added value workers, entrepreneurs and externalities and also a social one with an overall loss of talent. These threat were paired with the disappearance of retired workers and their retirement and early retirement pensions. This led to the risk of an urban decay process: reduced population, fewer services, more taxes and a less stable population.

As Pontes de García Rodríguez represents many of the points mentioned earlier:

- Here, city management focused on public-financed social and economic development from the public sector leading to positive externalities.

- However, in the absence of a collective vision for the city, public management failed to exploit multiple synergies while they were still viable to break away from the economic model of a monocultural industrial reality.

- There were clear indications of the public sector's passive attitude, whose intervention was non-existent, incomplete or misguided when challenged with a socioeconomic process of decline on such a large scale.

- Its most striking externalities (at least in regards to the priority that its internalisation must have) demonstrate a lack of expectations and negative self-image of their local population.

In fact, research was carried out by Escourido [3] through telephone conversations in a poll with a sample of 9,451 residents over 16 years of age, with a confidence level of 95% and 5% maximum sample error margin. Also interviewed was a panel of city representatives (Delphi, with total of 21 individuals). This study reveals low expectations with regards to the future of the city, based on the perception of the citizens.

This is considered a city with an ongoing deindustrialisation phase and a decadent image. Moreover there was a sensation of improvisation and inability from the public sector, as well a lack of social cohesion brought on by the duality between retired/early retired miners with high pensions and the rest of the population, affected by the loss of employment.

The internalisation (solution) of an externality closely related to the perceptions, desires and needs of the individuals has to take into account the need for public intervention aimed at reducing or eliminating transaction costs (information, property rights reallocation and negotiation). This must be carried out with the use of public and private collaboration mechanisms, marketing technology tools for cities which boost its self-image and strengthen a sense of belonging to the city. All of these measures are essential to develop subsequent socioeconomic recovery initiatives.

5. City marketing as a tool for public management to internalise the externalities of negative self-image. Case study of As Pontes de García Rodríguez (A Coruña).

Traditionally public management appeals to the political process for resource allocation; criteria include redistribution, equity, the common good or general interest. The level of focus depends on plurality, as well as administrative and executive fragmentation. Allocation, therefore, comes from legal actions and processes. In value creation, many factors are taken into account. These are variety and attachment, as well as transparency, zero discrimination, legality and social perceptions. This is especially true when the administered citizen acknowledges that the collective interest is not the sum total of all the individual interest, contrary to the way it is in the market when the word "client" is mentioned.

With economic efficiency goals -such as continuous improvement, resource utilisation and process based organised structure- and the citizen-client service, public management is more than a sovereign, coercive instrument. Instead, it has the end goal of providing services for society, such as identifying citizen needs and meeting them in an efficient and successful manner and with the highest possible quality. Marketing then becomes a technology that is adopted from the private sector for the public administration, to work as a detector and a way to measure the needs of the target market. Services can thus be developed to meet their needs and prepare for the upcoming changes. Every step is done in an efficient and successful way under the guidance of its vision, mission and values and goals.

This is the major advantage of city marketing when it is compared to other planning tools. The traditional focus of supply is supported by one of demand. This mechanism can spot and solve current needs by providing adequate services and infrastructure and future ones as well by developing new assets to improve living standards and keep competitive advantages in the target markets and establish communication in a way that a city brand is built. Moreover, the focus on demand allows for a more realistic and less resource intensive framework of action, with advantages derived from its efficiency.

One of the goals of city marketing is to improve and develop a positive city image [15]. Image has an enormous impact on city life as a "determining factor in urban development" [16]. This can become either a huge supporting factor or obstacle for the economic, cultural and social development of a city. It is considered a high importance "asset" in the development of city management [17].

This city image is built upon when defining its "vision". This is a must when planning projects based on demand (target market) and on identity resources. These are ideal for creating a guaranteed difference in the product [18]. The results will depend on the accuracy of the identity perception and the suitability of how the demand is adapted to each case. It is important to know beforehand the image of the city from the perspective of its locals and users: self-image is the starting point. Once the self-image is known, differences between the city's goals and those of its citizens, business and institutions can be asserted.

As mentioned above, if the starting point is a negative self-image of the city-no matter the origin of these externalities- the top priority should be to modify and restore the self-image in a positive way. This is also to restore a sense of belonging, along with local pride, to the city. It is the only way to make citizens take part in restoring the city and bring stakeholders back to plan for its future vision. With city marketing, the first objective is to define the city project based on the demand for its identified resources. This is the reason why a positive internal perception is needed in the city to minimise risks while elaborating future plans.

It is true that certain tangible elements of the city are required, like infrastructure or services, but these are not enough anymore. Nowadays intangible resources are what adds value and makes a

difference [19] when planning how to manage its capacity to attract people, investors, business, administration and tourism. In this sense, the top priority must be to fortify citizen-city relationships. The opinion of satisfied citizens about the city where they live creates a positive self-image of that city. Making sure every citizen has a certain living standard and services that meet their expectations and desires transforms them into active communicators of the city when they are outside of it. That is why it is vital to analyse the self-image of the city, to reinforce it or to correct a possible negative value.

When image policies are supported in a realistic vision of the city and the active participation of all its citizens is promoted by the administration, a brand image is created in a stronger, more efficient fashion. This is the opposite of a brand created solely by the local public servants. Image policy must strike a balance between image and quality of life, jointly coordinated by those responsible for that image policy, citizens and the municipal government.

Management of the image falls into the hands of its inhabitants, who are the inside target market, a vital task with the goal of fixing, improving and restoring a city's self-image. If the different inside target markets of the city are satisfied with or proud of the city they live in, they in turn help to make a better image and positive brand. If living standards, along with the social, cultural and economic environment, meet the diverse needs of citizens, business, and local institutions, each one of them will become a definer- or in marketing terms, a "preacher"- of the city brand.

To achieve this, internal communication must be suitably systematised so that information flows and positive effects are generated among all citizens and institutions within the city. This strengthens the values associated with pride and sense of belonging to or living in the city, innovation and change. Also, a positive self-image is the lever needed to involve, motivate and engage the stakeholders of the city with the city project. In the best case scenario, this project will be the product of participation, general agreement and internalisation.

This is a bottom-up approach, where city marketing expands the possibilities of strategic management and provides local authorities their participation in the hearings, discussions and decision making those groups and sections of the population who detect and pose in front of city problems [20], starting with the solution for the negative self-image.

In the case of As Pontes de García Rodríguez, there is access to different tools and resources to accomplish this internal communication. The city also has to confront one of the greatest challenges during the research phase. In particular, the panel of city representatives has to counter the existing negative connotations of the city's self-image. What the city is for its internal target markets and what they would like the city to be is an aspiration that must be present at all times. At the planning stage of the strategic city marketing, this must be true. These ideas must also inform the operational stage, when deciding which actions should be taken to meet the needs detected in the target markets of the city.

When the City Marketing Plan and all the measures mentioned in it are implemented, a general agreement must exist, with all public and private officers involved. This guarantees the continuity of the plan far beyond electoral horizons. The marketing applied to city planning means a long-term strategy aimed at creating new competitive advantages, one of which is the city brand itself. Also needed is the general support of all public and private agents within the city. To accomplish this, the contents of this plan must be agreed on by everyone and widely disseminated. Every strategic option planned for the city must have the approval of its stakeholders to avoid their rejection and with it the failure of its implementation.

The city and its planning are not described in terms of economic productivity and capacities. Instead, the focus should be on its attractiveness and the city's position in different global networks; this approach requires considerable collaborative and creative input. With the use of city marketing as a tool for local public management, interest (or what it is expressed as value) is the potential on a micro scale (the city) and as such it can be adapted to a macro scale (international flow system) to achieve a position as an attractive and well placed city in the global network. To do this, the city brand and its self-image and territorial coherence are essential, as is the sense of belonging, identity and planning to achieve it.

All without forgetting that the image of the city can be defined as organic, in continuous change [21]. If this image is negative (for example, after a process of long industrial decline), it requires a comprehensive application of city marketing (and not just advertising). This is because it is necessary to modify the self-image of the inhabitants to then be able to involve them in the transformation process that the city needs.

In this particular case (described in point 4), to improve self-image as well as recover self-esteem and sense of belonging among citizens, there is a series of measures that are to be carried out using city marketing as a tool for local public management [22] wich provides solutions for negative externalities. To correct these externalities, steps include [3]:

a) *Building a city network.* The objective is to create a configuration of networks allowing pressure groups to claim leadership positions and promote cities in a broad and ambitious way. It serves as a tool for implementing advanced technologies to gain access to a larger volume of information and exchange experiences at every city management level about supply and services, government, efficiency and management quality, environment and other topics. This involves integrating into and being part of an urban economic type network to find solutions to common problems and identify competitive advantages or share urban innovations. With cultural concerns, there can be promotion and defence of joint interests. The tourist domain is to promote tourism, culture, history and the environment, while the social content is about promoting values and cooperation. This is a communication and management tool for external projection, making the city more attractive and, above all, improving its self-image. The associations around an exemplary city and its attributes improve its perceived image both in- and outside that place.

b) *Promoting local heritage.* The historic, natural and industrial heritage in As Pontes de García Rodríguez is an element that connects past, present and future and projects the city's identity as a collective and/or a city. The footprints left behind by history, human interaction with nature and intense industrial activity in the area are a common link that can be used as a powerful tool to educate, build awareness and communicate in terms of an identity, an image and/or a brand. This natural and industrial heritage is used to preserve identity and self-image, while maintaining the continuity of the community and, at the same time, generating the right to use it and the collective duty of preserving it for the future.

A plan must be put in place to emphasise the value of its historic heritage, with archaeological sites from the Palaeolithic and Neolithic age, Roman period objects and bridge and the noble patronage of García Rodríguez de Valcárcel; its natural heritage, with is natural park in the Fragas do Eume region, lake, coal heaps, A Ribeira and Eume reservoirs and Caxado hills; and industrial heritage linked to mining operations and their industrial development. A process is needed to select what is of value and determine the next steps in incorporating collective identity and the actions that have to be taken to preserve and promote it.

It is important not to lose sight of the interpretation and value of these elements. This is done through logic, interests and perceptions because the planning must take into account general agreements about concepts, management models and value-deposed formulae.

c) *Reclaiming talent*. Building awareness to raise the level of talent inside the city lets it improve its self-image, as well as develop a sense of belonging and a collective pride. To shift perceptions and make the city more attractive to locals, events-such as conventions, shows and meetings- can be held to gather local people who excel professionally and have management positions in business or organisations.

d) *Mobilise and bond with the migrants*. As Pontes de García Rodríguez lived through the same migrations as the rest of Galicia during the end of the 19th century and beginning of the 20th (mostly migrants that went to South America) and again during the 50s and 60s (to South America and Centre of Europe). Later, the mining operations and power station made by Endesa, combined with the economic boom, allowed the vast majority of young people to have access to training and education at all levels showcasing a high percentage of citizens that reached higher education. Most of those students failed to find a job in their local city so they went to other areas.

At this point, the goal is to facilitate communication channels between "ponteses" (As Pontes locals) who are outside the city and keep and promote ties with the city. This measure can become an ideal economic, social and cultural tool among the members of the city and the institutions. These goals help to develop the city identity of the locals and their environment (family, social and economic); outside perception improves as the self-image of the locals gets better.

As shown, the courses of action proposed are aimed at public and private interests alike, during a time when people are solving public problems based on perceptions and personal feelings. This entails taking measures and holding discourse that brings satisfaction and meets needs. Also needed are a public worldview and narrative discourses (even political ones) that go deeper into the construction of the city from the public sector perspective.

With all of these points seen as tools for city marketing, the city is a space and an ensemble of meanings that are shaped and given function and purpose as the social agents relate to each other. They create a space to understand the diverse ways to express oneself, interpret symbols and produce meaning. This is where every audience is built from the collective senses, cooperation towards social change and encouragement of positive self-image and city pride.

Our perception of the world is almost always local, focusing, exclusively, on what is around us and our neighbourhoods, traditions and beliefs. At the same time, aspirations exist in all of our levels of society, from the lowest to the highest. Every one of us sends signals that express belonging to a tribe. If the big brands appeals to perceptions, needs, desires and feelings or hope, attractiveness, desirability, love, acceptance, and innovation, the image and brand of a city does not differ. As an individual and also as a citizen, I long for "what we had and lost" and "What we used to dream/be and never had/were" [23]. How we value this longing, evaluating and identifying the city and the image we have of it. To sum up, every community craves a strong sense of belonging and pride, which is the vehicle to any process of social change. This shows the importance of management from the local public perspective, the self-image of the city.

The aphorism "whatever the locals find unattractive outsiders will as well" is especially true in the case of As Pontes de García Rodríguez, as shown in the cited research. The appeal to city marketing from a public management perspective makes it possible, among other things, to act upon the immaterial values or intangibles of the city, in such a way that the self-image of the city shifts into something attractive for the residents. And it makes it possible, even in its application to small and medium-sized cities [24].

To do this, improving the reputation and citizen self-esteem, achieving a welcoming social environment open to innovation and lower the resistance to change (or adaptation) are the first steps needed to lay the foundations. Thus the main goal of public marketing is reached: enhancing the attractiveness of the city.

6. Conclussions

The importance of "good" management of the city is manifest, especially during a crisis. Traditional strategic planning focused on supply is not enough. An approach that focuses on demand is needed, along with the use of a city marketing technology.

In cities where the last recession has generated traumatic processes in terms of restructuring or the social and/or economic fabric has disappeared especially when they once enjoyed times of intense economic growth, one of the negative externalities with highest impact is the emergence of a negative self-image in the city. This in turns affects the city brand, expectations and decisions of its stakeholders and, with them, the future development of the city itself.

These externalities come from market failures and/or public sector mistakes and require public intervention to minimise or eliminate transaction costs, as Coase pointed out. Taking into account the nuances of subjectivism and perception management helps internalise (solve) the problem. This should be done by the citizens, and from there, act as a lever to promote social change and the future development of the city. To do this, city marketing used as a public urban management tool can be a solution for these externalities. When it is present, this must be done before implementing the rest of the City Marketing Plan, because the restoration of a positive self-image, pride and sense of belonging is a requirement to mobilise, integrate, involve and engage the stakeholders of the city to help the city reach the "vision" goal via the implementation of the rest of the plan

Every element is present in the case of As Pontes de García Rodríguez (Galicia, NW Spain). Public intervention made an industrial monoculture model that generated wealth and jobs for decades. When this model was exhausted and the state failed to intervene on a suitable scale, the city reached its present situation of decline. This decline is socioeconomic in nature, where the main externality is the generation of a negative self-image of the city, leading to a lack of expectations, depopulation and distrust in public management.

The use of city marketing as a tool for local and public management proposes as the top priority those actions that interiorise (solve) the negative self-image externalities in the city as the first step before anything else is done. This precedes any social or economic change in the city, and has to be done and planned for, taking into account the general approval and participation of the citizens who believe in the future and a vision for their city. This is a bottom-up approach.

It is important to note that marketing and the city are both in a continuous process of change and evolution. For this reason, the diagnosis stage could be carried out at a deeper level and report current perceptions of self-image through the use of new technologies: big/small data, the internet of things or social media.

At the same time, it is necessary to assess in which ways the immaterial aspects of the city gain importance over the material and to see how they flow, influencing the definition of new shapes and forms of city planning and management. This can be done by integrating new methodologies of citizenship participation and social creativity. It would be useful to contrast "city branding" from the point of view of public policy supported by the strategic planning of the city to enhance its capacities and unfold its competitive advantages.

References

[1] R. Camagni, Economia Urbana, Antoni Bosch Editor, 2005.
[2] D. Harvey, The Urban Experience, Basil Blackwell, 1989.
[3] M. Escourido, "El marketing de ciudades como herramienta de gestión pública local: una aplicación al caso de As Pontes de García Rodríguez (A Coruña)", Ph.D Thesis, A Coruña University, 2017.
[4] C. Tiebout, "A Pure Theory of Local Expediture", The Journal of Political Economy, 64 (5), 416-424, 1956.
[5] P. Kotler, D. Haider, I. Rein, Mercadotecnia de localidades, Ed. Diana, 1994.
[6] D. Weimer, A. Vining, Policy Analysis and Concepts, Prentice Hall, 2004.
[7] A. C. Pigou, The Economics of Welfare, MacMillan, 1932.
[8] R. H. Coase, La empresa, el mercado y la ley, Alianza Ed., 1994.
[9] R. Cooter, T. Ulen, Law and Economics, Harper Collins, 1988.
[10] J. M. Aguilera, "El análisis económico del derecho de propiedad: una aproximación", Anuario Jurídico y Económico Escurialense, XLII, 117-128, 2009.
[11] F. A. Hayek, The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology, Kessinger Legacy Reprints, 1952.
[12] L. M. Kirzner, "Competition and Entrepreneurship", University of Chicago Press, 1973.
[13] R. Thaler, Todo lo que he aprendido con la psicología económica: el encuentro entre la economía y la psicología, y sus implicaciones para los individuos, Ed. Deusto, 2016.
[14] S. C. Littlechild, "El problema del costo social", Revista Libertas 12, Instituto Universitario Eseade, 1990.
[15] M. Escourido, "La marca ciudad: la importancia de su valor", Revista ACEF-UDIMA, 2013.
[16] R. Friedmann, "Marketing urbano", Contactar, 3, 1999.
[17] M. A. Akbar et al., "An empirical Study Investigation of Task Allocation Process Barriers in the Context of Offshore Software Development Outsourcing: An Organization Size Based Analysis," Int. J. Comput. Digit. Syst., vol. 8, no. 04, pp. 343–350, 2019.
[18] A. Precedo, M. Escourido, Marketing de ciudades y territorios: una visión renovada, Dextra Ed., 2017.
[19] S. Rivero, "Importancia de los intangibles como motores del crecimiento metropolitano", Curso Gestión Estratégica de Ciudades, Asociación Bilbao Metrópoli 30, 2001.
[20] A. Kiseleva, "Marketing Strategy of the Territory in the Investment Policy of the City", Strategic Management, 22 (1), 22-29, 2017.
[21] G. Hankinson, "Managing Destination Brands: establishing a Theoretical Foundation", Journal of Marketing Management, 25 (1-2), 97-115, 2009.
[22] A. Martin, "Marketing a Municipality as a Place. A Study of Three Municipalities in the north of Sweden", Bachelor Tesis, Lulea University of Technology, 2009.
[23] M. Lindstrom, Small Data, Ed. Deusto, 2016.
[24] N. Dril, A. Galkin, N. Bibik, "Applying City Marketing as a Tool to support Sustainable Development is Samal Cities: case estudy in Ukraine", 2nd International Conference "Green Cities – Green Logistics for Greener Cities", Szczecin, Poland, 2016.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 6, 115-125 (2019)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Learning Literary Style End-to-end with Artificial Neural Networks

Ivan P. Yamshchikov[*,1], Alexey Tikhonov[2]

[1]Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

[2]Yandex, Berlin, Germany

ARTICLE INFO

Article history:
Received: 20 June, 2019
Accepted: 26 October, 2019
Online: 28 November, 2019

Keywords:
Stylized text generation
Poetry generation
Artificial neural networks

ABSTRACT

This paper addresses the generation of stylized texts in a multilingual setup. A long short-term memory (LSTM) language model with extended phonetic and semantic embeddings is shown to capture poetic style when trained end-to-end without any expert knowledge. Phonetics seems to have a comparable contribution to the overall model performance as the information on the target author. The quality of the generated texts is estimated through bilingual evaluation understudy (BLEU), a new cross-entropy based metric, and a survey of human peers. When a style of target author is recognized by the humans, they do not seem to distinguish generated texts and originals.

1 Introduction

This paper is an extension of work presented initially in [1] enhanced with the reasoning and experiments described in [2] and [3].

Generation of authentic, stylistically expressive texts is still challenging despite numerous recent advancement in machine-generation of texts. The ability to generate texts that feel personal and expressive opens numerous possibilities for industrial and scientific applications, see [4] or [5]. For example, the generation of stylized texts could improve user experience in human-machine interfaces. When talking about interactions between a human and a machine, one tends to speak in terms of user experience or interface instead of communication. This is partly due to the fact that predictability in a narrow context was for many years (and still is) one of the main aspects of human-machine interfaces [6], [7]. Communication, on the other hand, naturally implies a certain degree of surprise and context shifts on the way [8]–[10]. We expect understandable and reproducible behavior when interacting with a machine, yet regularly allow other humans to be more versatile and obscure when communicating with us. Yet modern technologies that comply with these expectations tend to be perceived as flat and dull. Indeed, personal assistants such as Alexa or Siri have limited anthropomorphic features. However, they can not be tailored to the needs and preferences of every end-user. Other AI-powered solutions tend to have no anthropomorphism at all, yet humans voluntarily interact with them. One such example would be a game of advanced chess. In advanced chess, two players accompanied by two AIs play against one another. The concept of such a game was proposed in the 1970s [11], but the history of advanced chess started in 1998, just one year after the historic game between Garry Kasparov and Deep Blue. It represented a paradigm shift, from the twentieth-century 'humans against the machines' to the twenty-first-century 'humans with the machines.' Stylistically expressive text generation can facilitate and speed up this shift.

Modern generative models are typically trained on massive corpora of texts. These corpora consist of various texts created by different people in diverse genres and forms. Standard generative models do not incorporate the information on authorship in any manner and therefore learn to generate texts that are some sort of 'average' across the training dataset. Such texts could hardly be perceived as human-written. Indeed, there is an intuitive belief that each human has a recognizable writing style. This intuition, however, is not backed up by any quantitative definition of a literary style. Experts can have an extensive discussion on what distinguishes a text written by Edgar Allan Poe from a text written by Walt Whitman. However, as far as we are concerned, these expert opinions can rarely be quantified in the context of a rigorously defined natural language processing (NLP) task. That is the reason why this paper addresses the problem of author styl-

*Corresponding author: Ivan P. Yamshchikov, Inselstrasse 22, 04103, Leipzig, Germany, ivan@yamshchikov.info

ized text generation in an end-to-end setup. We propose a model that generates texts resembling the writing style of a particular author without any external information on stylistic aspects of the author's writing. We also restrict the scope of this paper to poetry rather than prose due to several reasons. First, the same machine learning model trained on a smaller dataset tends to show weaker performance. Naturally, for a given amount of authors, a poetic corpus would be inevitably smaller in terms of the number of available texts than a corpus that includes prose. This makes the problem more challenging. However, one can argue that if a proposed method manages to reproduce style when trained on a relatively small corpus of poems, it would be applicable to a bigger dataset as well. Second, poetry is often regarded as a more stylistically expressive form of literary art. This gives hope that even with relatively short chunks of generated texts, one could easily distinguish the style of a target author. One of the methods to estimate the performance of the generative model is to measure how often humans attribute the generated texts to the target authors. The expressive nature of poetry makes such assessment easier for the peers and, therefore, facilitates the experiments. Both these factors make poetry generation a more challenging and a more exciting task than the generation of prose. Nevertheless, we have every reason to believe that a model that successfully generates stylized poetry would also successfully mimic the style of a prosaic text.

This paper has four significant contributions: (1) formalization of stylized text generation framework; (2) a new sample cross-entropy metric that could be used to measure the quality of the stylization; (3) a long short-term memory artificial neural network with extended phonetic and semantic embeddings that generates stylized texts that are quantified both subjectively through a survey and objectively with sample cross-entropy and BLEU metrics; (4) the proposed approach is applicable to a multilingual setting. Most importantly, we show that modern generative neural networks can learn and mimic the author's style in an end-to-end manner without any external expert knowledge.

2 Related work

Almost a century ago, [12] suggested that computers can generate poetry algorithmically. In [13], one can find a detailed taxonomy of generative poetry techniques. However, to work with stylized poetry generation, one has to resolve the ambiguity of the literary style. In a rapidly developing field of textual style-transfer, one can see various definitions of literary style that often contradict one another [14]. Indeed, different aspects of literary style could includ a sentiment of a text (see [15] or [16]), it's politeness [17] or a so-called style of the time (see [18]). The style of the time aspect is also addressed by [19] and by [20]. A paper by [21] generalizes these ideas and proposes to measure every stylistic aspect with a dedicated external classifier. Yet the problem of style transfer differs from the stylized text generation significantly: a human-written input could be used to control the saliency of the output [22]. This might improve the resulting quality of the texts that are generated. In the meantime, such input is not available for the generative model, and it has to generate stylized texts from scratch. This makes our problem setup, similar to [23]. This paper shows that modern artificial neural networks can learn the literary style of an author end-to-end without any predefined expert knowledge. This is a meaningful empirical result in itself since it serves as an initial proof-of-concept for further research of stylized text generation. In this paper, we focus on RNN-based generative models. Let us briefly mention several contributions in that area that are relevant for further discussion.

There are a few papers that decompose content and semantics of the input, see [24]–[29]. Yet none of them works with the literary style of the generated texts. In [30] the authors developed a persona-based model that was to be consistent in terms of response generation. There are works that specifically address Chinese classical poetry generation, for example [31]–[35]. However, the contributions in the area of generative poetry in languages other than Chinese or in a multilingual setting are relatively rare. In [36], an algorithm generates a poem in line with a topic given by the user, and in [37], the authors generate stylized rap lyrics with LSTM trained on a rap poetry corpus. These contributions are all based on the idea that when trained on a corpus of similar documents, an artificial neural network could infer certain properties of such text and reproduce them in generative mode. We suggest to make one step further and develop a model that could be trained on a broad dataset of various documents yet generate texts that resemble various subsets of the training dataset.

3 Generation of stylized texts

Consider a corpus $C = \{T_i\}_{i=0}^{M}$ of M texts written in a given language. Every text of length l is a sequence $T_i = (w_j)_{j=0}^{l}$ where words (denoted here as w_j) are drawn from a vocabulary set $V = \{w_j\}_{j=1}^{L}$, where L is the size of a given vocabulary. One can try to vary the length of T_i in training. For practical reasons we further assume that T_i are lines of poems. Sine these are the text in natural language there is additional structure over the set V and certain principles according to which sequences T_i are formed. However, this information is only partially observable.

In a standard problem of text generation, a language model predicts the next word w_k using a conditional probability $P(w_k|(w_i)_{i=0}^{k-1})$. Neural networks are widely used for language modeling since [38], see also [39] and [40]. In particular, this family of algorithms allows to avoid the dimensionality curse of a classical language model. One typically would like to obtain a mapping $Y : (C, \mathbb{R}^m, F) \to \mathbb{R}^d$ and then train a model such that $G(C) : \mathbb{R}^d \to \{T_i^G\}$.

Usually one tries to use additional observable information in order to improve the general performance of the model, e.g. [41]. For example, if there is a performance metric D (such as BLEU, F1, etc.), one usually minimizes $D(\{T_i\}, \{T_i^G\})$, where $\{T_i\}$ is a randomized sample of C. Let us find a stylization

model $G(C|S)$ that takes into consideration a subset S of continuos and categorial variables out of $(\mathbb{R}^m, F)$ and a metric D in such a way that

$$G(C|S): \begin{cases} (C, \mathbb{R}^m, F) \to \{T_i^G\} \\ \{T_i^G|S\} \sim \{T_i|S\} \text{ w.r.t. } D \end{cases} \quad (1)$$

A stark difference in our approach is that we train our model using all information available to us, i.e., $(C, \mathbb{R}^m, F)$. However, we do not measure its overall performance, but rather test it on a specific domain S. This core idea is, in some sense, similar to the principles of one-shot learning, see [42], and, generally, transfer learning, see [43], and author-attribution methods, see [44]. The model has access to information about a broader domain of data. This information happens to exogenous to the problem in its' narrow formulation. For a given domain S, it can seem irrelevant, yet it can significantly boost the results of the model. There are several advantages associated with this approach. First of all, such models naturally imply ample opportunities for customization. If one wants some additional control, the parameters of the model one can include them into S. The output $\{T_i^G|S\}$ will resemble original texts $\{T_i|S\}$ that satisfy S conditions. That property makes such methods applicable to personalized interfaces. At the same time, one would expect that due to the 'umbrella' structure in which $G(C|S)$ is trained on the whole corpus $(C, \mathbb{R}^m, F)$ the model would outperform various smaller models trained on different subsamples of C. Artificial neural networks are known to generalize well. This allows to speculate that system he system that uses less information for training would show inferior performance when compared with a system that is trained on the whole corpus C.

Further we work with an artificial neural network that uses the name of an author as such condition S. Such models when trained on a sufficiently large dataset, generates lyrics similar to the text written by a target author. The quality of the stylization is assessed both subjectively (based on a survey of respondents) and objectively (in terms of BLEU and a sample cross-entropy that we define further). The model has been trained with English and Russian datasets. There are no obstacles to implement this model in other languages. Before we describe our model in detail, let us briefly address one of the issues of poetry generation that rarely gets enough attention from researchers, namely, the phonetics of the lyrics.

4 Importance of phonetics

The structure of the poem could be different across different languages. For example, Chinese poems have highly specific and distinct structures, see [35], whereas some American poetry of the twentieth century or so-called white poems in Russian hardly have any definite structure at all. The structure of poems can depend on various factors. These factors are primarily phonetic. In the broadest sense, a sequence of tones in a Chinese quatrain, a structure of a Persian ruba'i, rhymes in a classical western sonnet, or a structure within rap bars could be expressed as a set of phonetic rules based on a particular understanding of expressiveness and euphony shared across a given culture. Some cultures and styles also have particular semantic limitations or 'standards,' for example, 'centrality' of specific topics in classical Japanese poetry, see [45]. We do not make attempts to address high-level semantic structure. However, one can add some pseudo-semantic rules to our model, say via some mechanism in line with [36] or [32]. Instead, we suggest focusing on syntax and phonetics.

The importance of phonetics in poetical texts was broadly discussed among Russian futuristic poets, see [46]. Several Russian linguistic circles and art groups (particularly OPOJAZ) in the first quarter of 20th century were actively discussing the concept of the abstruse language, see [47], also stressing that the form of a poem, and especially its acoustic structure, is a number one priority for the future of literature. In their recent paper, [48] have challenged the broadly accepted idea that sound and meaning are not interdependent: unrelated languages very often use (or avoid) the same sounds for specific referents. [49] backs this idea up, showing that modern word embeddings capture such dependencies to a certain respect. In line with these ideas, one of the critical features of the model that we discuss below is its concatenated embedding that contains information on the phonetics of every word preprocessed by a bi-directional LSTM alongside with its vectorized semantic representation.

5 Model

Let us look at an LSTM-based language model that predicts the w_{n+1} word based on $w_1, ..., w_n$ previous inputs and other macro parameters of the sequence. Typically one passes the needed parameter to the network by writing it in its initial state. However, as the generated sequence gets longer, the network often 'forgets' the parameters of the document. We want to develop a model to address the problem given in (1). To do that, we support our model at every step with the embeddings of the document that is currently being analyzed (these are variables from the set S on which we want to condition our model during the generation process, such as the name of a document or its author). This idea makes our approach different from a classical word-based LSTM. It was used in [50] for stylized music generation. Figure 1 shows a schematic picture of the model; document information projections are highlighted with blue and white arrows.

We used an LSTM with 1152-dimensional input and 512-dimensional projection of a concatenated author and document embeddings. A concatenated word representation is shown schematically in Figure 2. At every step and use the word embeddings, which are generally used for a task of this type. However, we believe that word embeddings are not enough: the char-based information might allow for some useful grammar rules to be learned, and, as we discussed in the previous section, phonetics plays a crucial role in the euphony of the final output. To include this information, we suggest using two bidirectional LSTMs with a 128-dimensional vector as a final state. One of the LSTMs works with a char-representation of

the word whereas another uses phonemes of the International Phonetic Alphabet[1], employing heuristics to transcribe words into phonemes. Before training the model, we obtained a vocabulary V for the training dataset, then applied transcribing heuristics to it. The result was a phoneme-based vocabulary V_p. Then, in the training phase, the char-based bidirectional LSTM was reading words from the initial vocabulary $w_j \in V$, and the phoneme-based LSTM was reading words from the heuristically transcribed vocabulary $w_j^{'} \in V_p$. [51] proposes a somewhat similar idea, but with convolutional neural networks rather than with LSTMs. The bidirectional LSTM approach is new to our knowledge.

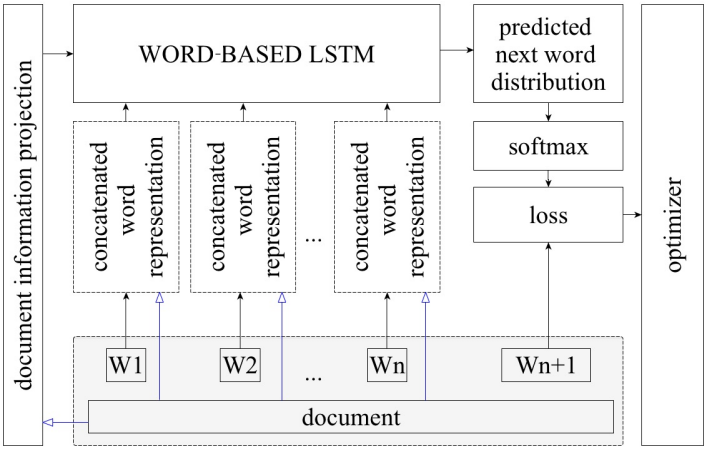


Figure 1: The scheme of the language model. Document information projections are highlighted with blue and white arrows. The projections on a state space of the corresponding dimension is achieved with simple matrix multiplication of document embeddings.

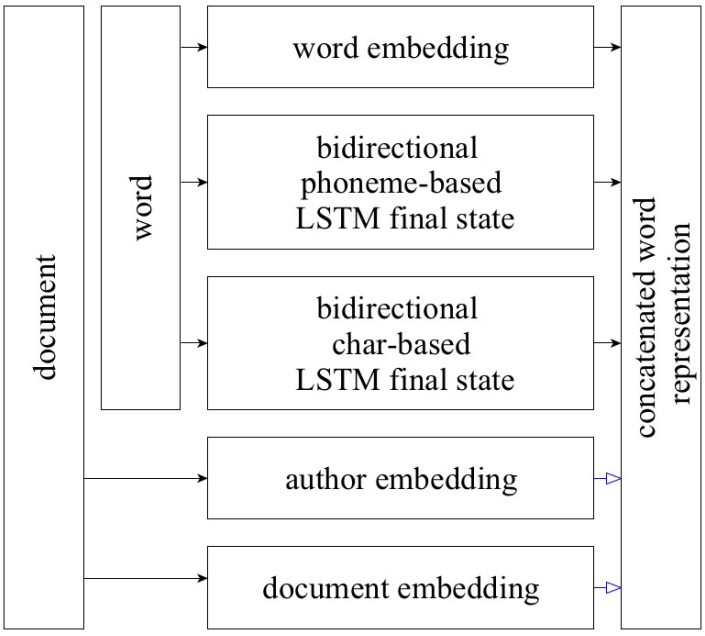


Figure 2: Concatenated word representation.

[1]https://en.wikipedia.org/wiki/International_Phonetic_Alphabet

In Section 7, we describe a series of objective and subjective tests that we ran across a generated output $\{T_i|S\}$, but first, let us briefly describe the datasets used for training.

6 Datasets

Two proprietary datasets of English and Russian poetry were used for training. Every character was transferred to a lower case, and punctuation was deleted. No other preprocessing was made. The overview of the sizes of the datasets can be found in Table 1. This allowed having approximately 330 000 verses in train dataset and another 10 000 verses forming a test dataset for Russian poetry. For English poetry train data consisted of 360 000 verses with approximately 40 000 verses forming the test data.

Table 1: Parameters of the training datasets.

	N. of documents	Size of vocab.	N. of authors	Size
English	110000	165000	19000	150 Mb
Russian	330000	400000	1700	140 Mb

Table 2: Number of words in the training datasets for human-peer tested lyrics.

	N. of words		N. of words
Shakespeare	10 218	Pushkin	226 001
Carroll	19 632	Esenin	73 070
Marley	22 504	Letov	29 766
MUSE	7 031	Zemfira	23 099

During the training phase, the beginning and end of every text T_i were tokenized. In the generation phase, the network is conditioned on values of document parameters S and is initialized with a special 'start' token. As we stated previously, in the most general setup, the set of variables S on which we want to condition our generative model $G(S)$ can vary. Here we use only one categorical variable - the name of the author to test the proposed mechanism for the stylized text generation. We trained the model for Russian (Alexander Pushkin, Sergey Esenin, Joseph Brodsky, Egor Letov and Zemfira Ramazanova) and for English (running tests on lyrics of William Shakespeare, Edgar Allan Poe, Oscar Wilde, Lewis Carroll and Bob Marley as well as lyrics of the American band Nirvana and UK band Muse). Table 2 shows that there were far more authors in the dataset. For the experiments, we chose more prominent authors who are known for their poetic styles. A fluent speaker of the target language could easier identify these authors. The model produces results of comparable quality for both languages, so to make this paper shorter, we further address generative poems in English only and provide the experimental results for Russian in the Appendix. We do not see any difficulties that could prevent the implementation of the proposed model in other languages. As long as there

is a sufficiently large corpus C and a phonetically transcribed vocabulary V_p, one could use such a method.

The model produces results of comparable quality for both languages, so to make this paper shorter, we further address generative poems in English only and provide the experimental results for Russian in the Appendix.

Table 3 shows some examples of the generated stylized poetry. The model seems to capture the syntactic characteristics of the author (look at the double negation in the Neuro-Marley lyrics, for example). The vocabulary that is characteristic for a target author alsoo seems to be captured. 'Darkness,' 'burden,' 'fears' could be subjectively associated with the gothic lyrics of Poe. At the same time, 'fun,' 'sunshine,' 'fighting every rule' could be associated with positive yet rebellious reggae music. Author-specific vocabulary can technically imply specific phonetics that characterizes a given author; however, this implication is not self-evident, and generally speaking, does not have to hold. As we demonstrate later, phonetics does, indeed, contribute to the author's stylization significantly.

7 Experiments and evaluation

There is a variety of metrics that could be used as D, for which we could estimate the quality of the generated texts. A standard approach for a comparison of two generative models is to measure cross-entropy loss at certain checkpoints. However, as [52] writes: "There can be significant differences in final performance across checkpoints with similar validation losses." In the case of the provided model, a cross-entropy does not seem to give any meaningful information. To quantitatively estimate the final model $G(C|S)$, we carried out an ablation experiment. We used a plain vanilla LSTM without word-by-word document information support that was only using classic word embeddings as a baseline. Another model to compare with included document information support but didn't have bidirectional LSTMs for phonemes and characters. All the models have shown comparable values of cross-entropy loss after an equal amount of epochs. It seems that the additional structures are not facilitating learning but are also not hindering it. Below we describe two automated metrics that are natural for the problems of such type, namely, the BLEU and the sample cross entropy. We demonstrate that the model captures styles of different authors and that instead of developing several different models $\{G(S_i)\}$ trained on corpora corresponding to different authors one can use one model $G|S$ and change S depending on the style that one wants to reproduce.

7.1 Sample cross entropy

Cross entropy seems to be one of the most natural theoretic-informational metrics to estimate the similarity of different discrete sequences. To distinguish it from the cross-entropy loss, we further call it the sample cross entropy. We calculate it as follows. One samples several subsets with the same length (in words) from the original author texts. The sample is drawn so that there is a comparable number of unique texts for each author within the sample. Then one splits the texts of a given author A_i into two random groups and calculates the pairwise[2]cross-entropy between original texts of the author A_i and texts generated by the model $\{T_i^G|A_i\}$. The cross-entropy between the obtained sets of texts was calculated with MITML, see [53], as follows. For every sample written by the author, one built a standard 3-gram based language model with standard MITML smoothing. A shared vocabulary across all samples was calculated. Then the perplexity was computed by applying the language models based on the author-written texts to the original and generated texts. Though both values have a similar meaning, we suggest applying logarithm to the obtained value to get the cross-entropy instead of the perplexity. In Table 4 one can see the results of these calculations. Analogous results for Russian can be found in Appendix in Table 10. The model clearly captures the style of the time mentioned earlier alongside the individual styles. Generated texts stylized for the authors from a similar time period tend to demonstrate lower sample cross-entropy with human-written texts written close to that time.

The model captures the author's writing style better if the sample cross-entropy between the texts written by the given author and the texts generated by the model is lower. We also provide cross-entropy between random samples from the texts of the same author to demonstrate how self-similar are the human-written texts. The majority of texts in the training dataset were the texts from the 20th century. One can see that the model 'perceives' texts of Edgar Allan Poe or William Shakespeare to be closer to the lyrics of Oscar Wilde and Lewis Carrol than to the samples of the original texts. At the same time, Poe and Shakespeare are also reasonably well approximated by the model. To give a reference, we also provide cross-entropies between the texts sampled randomly out of the vocabulary as well as the texts obtained through a weighted average sampling and the human-written texts.

7.2 BLEU

BLEU is a metric estimating the correspondence between a human-written text and a machine's output. It is typically used to assess the quality of of machine translation. We suggest to adopt it for the task of stylized text generation in the following way: a random starting line out of the human-written poems is used to initialize the generation. Generative model 'finishes' the poem generating thee ending lines of the quatrain. Then one calculates BLEU between lines generated by the model and three actual lines that finished the human-written quatrain.

First of all, one can see that it makes sense to train the model on the whole dataset and then restrict it to the chosen author rather than train it on the target author exclusively. This is shown in Table 5.

[2]Hence 'sample' in the name of the metric.

Table 3: Examples of the generated stylized quatrains. The punctuation is omitted since it was omitted in the training dataset.

Neuro-Poe	Neuro-Marley
her beautiful eyes were bright	don t you know you ain t no fool
this day is a burden of tears	you r gonna make some fun
the darkness of the night	but she s fighting every rule
our dreams of hope and fears	ain t no sunshine when she s gone

Table 4: Sample cross entropy between generated texts $\{T_i^G|A_i\}$ and actual texts for different authors. The two smallest values in each row are marked with * and ** and a bold typeface. The sample cross entropy between random samples from the texts of the target author and randomly generated sequences of words (uniform and weighted respectively) as well as other samples written by the same author (denoted as SELF) are shown for reference.

Model $G(A_i)$/ author	Shakespeare	Poe	Carroll	Wilde	Marley	Nirvana	MUSE
Neuro-Shakespeare	**19.0****	21.6	**18.5***	19.9	21.8	22.0	22.4
Neuro-Poe	22.0	**20.4****	21.2	**19.0***	26.0	25.4	26.0
Neuro-Carroll	22.2	23.6	**18.9***	22.5	22.4	**21.8****	23.8
Neuro-Wilde	21.2	20.9	**20.5****	**18.4***	24.5	24.8	26.4
Neuro-Marley	24.1	26.5	22.0	27.0	**15.5***	**15.7****	16.0
Neuro-Nirvana	23.7	26.2	20.0	26.6	19.3	**18.3***	**19.1****
Neuro-MUSE	21.1	23.9	18.5	23.4	17.4	**16.0****	**14.6***
Uniform Random	103.1	103.0	103.0	103.0	103.5	103.3	103.6
Weighted Random	68.6	68.8	67.4	68.5	68.5	68.0	68.0
SELF	23.4	21.8	25.1	27.3	20.8	17.8	13.3

Table 5: BLEU for the full model trained on one particular author dataset, $G(S)$, and on the whole dataset, $G(C|S)$, calculated on the chosen author validation dataset and on the validation dataset that includes a variety of authors. The results may vary across authors depending on the relative sizes of S and C but the general picture does not change.

Model $G(A_i)$	Chosen author S	Validation dataset
$G(S)$	33.0%	19.0%
$G(C\|S)$	37.3%(+13%)	37.6%(+98%)

The model $G(S)$ is trained on texts of a particular domain S. In this case the domain corresponds to a particular author. One can see that $G(C|S)$ when tested on the lyrics of the chosen author outperforms $G(S)$. Moreover, $G(C|S)$ performs almost twice as good as $G(S)$ on the validation dataset comprised out of the texts of other authors.

Table 6 summarizes the results of the ablation experiments. One can see BLEU calculated on the validation dataset for the plain vanilla LSTM, LSTM extended with author information but without bidirectional LSTMs for phonemes and characters, and the full model. The uniform random and weighted random give baselines to compare the model.

Table 6: BLEU for uniform and weighted random random sampling, vanilla LSTM, LSTM with author embeddings but without phonetics, and for the full model. Phonetics is estimated to be almost as important for the task of stylization as the information on the target author.

Model $G(A_i)$	BLEU
Uniform Random	0.35%
Weighted Random	24.7%
Vanilla LSTM	29.0%
Author LSTM	29.3% (+1% to vanilla LSTM)
Full model	29.5% (+1.7% to vanilla LSTM)

Table 6 shows that extended phonetic embeddings play a significant role in the overall quality of the generated stylized output. It is essential to mention that phonetics is an inherent characteristic of an author and the training dataset. Humans do not have qualitative insights into phonetics of Wilde or Cobain, yet the information on it turns out to be necessary for the style attribution.

7.3 Survey data

For a survey we sampled quatrains from William Shakespeare, Lewis Carroll, Bob Marley, and MUSE band. Each sampled quatrain was accompanied by a quatrain generated by the model conditioned on the same author. One hundred and forty fluent English-speakers were asked to read all sixteen quatrains in a randomized order and choose one option out of five offered for each quatrain. They had to decide if the author of a given verse was William Shakespeare, Lewis Carroll, Bob Marley, MUSE, or an Artificial Neural Network. Table 7 shows the summary of the obtained results. For analogous results in Russian see in Table 11. A more detailed description of the methodology could be found in the Appendix.The generated texts were human-filtered for mistakes, such as demonstrated in Table 8. This stands to reason since clear mistakes would inevitably allow human peers to detect a generated text. The automated metrics provided above were estimated on the whole sample of generated texts without any human-filtering.

Table 7 shows that the model manages to stylize several given authors. The participants recognized Shakespeare more than 46% of the time (almost 2.5 times more often than compared with a random choice). They did slightly worse in recognizing Bob Marley (40% of cases) and MUSE (39% of

Table 7: Results of a survey with 140 respondents. Shares of each out of 5 different answers given by people when reading an exempt of a poetic text by the author stated in the first column. The two biggest values in each row are marked with * and ** and a bold typeface.

Model $G(A_i)$ or author	Shakespeare	Carroll	Marley	MUSE	LSTM
Neuro-Shakespeare	**0.37***	0.04	0.05	0.14	**0.3****
Shakespeare	**0.46***	0.05	0.04	0.07	**0.3****
Neuro-Carroll	0.02	0.07	**0.26****	0.18	**0.41***
Carroll	0.05	**0.2****	0.14	0.11	**0.32***
Neuro-Marley	0.02	0.01	**0.47***	0.2	**0.29****
Marley	0.15	0.05	**0.4***	0.1	**0.24****
Neuro-MUSE	0.09	0	0.12	**0.34****	**0.39***
MUSE	0.03	0.05	**0.28****	**0.39***	0.2

cases, 2 times higher than a random choice). It seems that the human-written quatrains were recognizable. The participants were also fluent enough in the target language to attribute given texts to the correct author. However, people believed that the generated text was actually written by a target author in 37% of cases for Neuro-Shakespeare, 47% for Neuro-Marley, and 34% for Neuro-MUSE, respectively. Lewis Carroll was less recognizable. His texts were only recognized in 20% of cases. The underperformance of the model can be explained with difficulty experienced by the participants in determining Carroll's authorship. If humans have a hard time recognizing the author's style, it is hard to expect an excellent performance of a model on such a target author. Combining the results in Table 6 with the results of the survey shown in Table 7, one could conclude that the phonetic structure of lyrics has an impact on the correct author attribution of the stylized content. This impact is usually not acknowledged by a human reader but is highlighted with the proposed experiment.

7.4 Discussion

Here we would like to share some of the insights into the nature of computational creativity. First, the stylized generation is thought-provoking since the model learns similarities between various authors in the dataset and extrapolates if there is a deficit of information. For example, one of the Nirvana-stylized lines ran:

a god who's always welcome to iraq

Kurt Cobain committed suicide long before the start of the Iraq campaign. However, since the model was trained on a massive corpus of poems (and not only on Nirvana ones), this line emerged among the generated lyrics. It also fits into the broader context of a song due to the correct rhythmic structure. When filtering the network's output, we found this line exceptionally interesting. It was meaningful to us because it resonated with our intuition that, if he were alive, Cobain would have spoken about the Iraq war in his lyrics. We also found it stylistically similar to Cobain lyrics and emotionally provocative. Formally speaking, all of these characteristics are human-attributed, and yet they can emerge out of explorative human-algorithm interactions. One could reason that the topic of Iraq occurred in other songs by grange bans, and the model extrapolated the saliency of the topic to the Cobain lyrics.

Second, creativity is serendipitous and highly depends on the perception of the viewer. In the Appendix of this paper, we provide examples of typical mistakes. One such mistake that is characteristic for the state of the art artificial neural network is an abrupt end of a line. As we listed the examples of such behavior, we came across the following line:

at night i lay waiting for a

However, one of our colleagues pointed out that this line represents a perfect one-line poem; an abrupt ending only enhances the feeling of longing. We have decided to submit this poem among others into the journal of modern poetry, and it was published[3]. We find both these insights meaningful for further research on computational creativity.

Over the last several years, we have seen several attempts at employing AI as a tool for creators. Using the approach proposed in [54], some visual artists applied style transfer algorithms to various pictures, creating new and unique experiences. In a postmodern cultural context, style transfer can be a conceptual tool that is able to convey the thought of an artist in a more expressive and meaningful way, for example, see [55].

8 Conclusion

This paper addresses the problem of author-stylized text generation. We show that an extension of a language model supported by the document meta-information at every step is apt for such tasks. Large concatenated embeddings that consist of word embedding, a phoneme-based bidirectional LSTM final state, and a char-based bidirectional LSTM final state are shown to facilitate end-to-end learning of author poetic style. Moreover, extending word embeddings with phonetic information has a comparable impact on the BLEU of the generative model as the information on the authors of the text. A cross-entropy method to estimate the quality of stylization is proposed. This model is shown to perform in Russian and English, and there seem to be no obvious reasons for which one could not implement it in other languages. It was shown

[3] https://nokturno.fi/poem/on-author-stylized-neural-network-poetry

that the texts generated by the model tend to be closer to the texts of the target author than the text generated by a plain vanilla LSTM both in terms of the cross-sample entropy and BLEU. It was also shown that, when faced with an author with a recognizable style (an author who is recognized approximately two times more frequently than at random), humans mistakenly recognize the output of the proposed generative model for the target author as often as they correctly attribute original texts to the author in question.

References

[1] Alexey Tikhonov and Ivan P. Yamshchikov. Guess who? multilingual approach for the automated generation of author-stylized poetry. In IEEE Spoken Language Technology Workshop (SLT), pages 787–794, 2018.

[2] Alexey Tikhonov and Ivan P. Yamshchikov. Sounds wilde. phonetically extended embeddings for author-stylized poetry generation. In Proceedings of the Fifteenth Workshop on Computational Research in Phonetics, Phonology, and Morphology, pages 117–124), 2018.

[3] Ivan P. Yamshchikov and Alexey Tikhonov. I feel you: What makes algorithmic experience personal? In Proceedings of EVA Copenhagen 2018 - Politics of the Machines - Art and After, 2018.

[4] Daniel Livingstone. Turing's test and believable ai in games. Computers in Entertainment, 4(1):6, 2006.

[5] Alan Dix. Human-like computing and human—computer interaction. In Proceedings of the 30th International BCS Human Computer Interaction Conference: Fusion!, page 52, 2016.

[6] J. Weizenbaum. Eliza—a computer program for the study of natural language communication between man and machine. Communications of the ACM, 9(1):36–45, 1966.

[7] S. K. Card, T.P Moran, and A. Newell. The psychology of human-computer interaction. CRC Press, 1983.

[8] Y. Ohsawa. Modeling the process of chance discovery. Chance discovery, pages 2–15, 2003.

[9] W. R. Saab, N. abd Joolingen and B. H. Hout-Wolters. Eliza—a computer program for the study of natural language communication between man and machine. British Journal of Educational Psychology, 75(4):603–621, 2005.

[10] A. Abe. Curation and communication in chance discovery. In Proceedings of the 6th International Workshop on Chance Discovery (IWCD6) in IJCAI, 2011.

[11] D. Michie. Programmer's gambit. New Scientist, 17th of August:329–332, 1972.

[12] Jon Wheatley. The computer as poet. Journal of Mathematics and the Arts, 72(1):105, 1965.

[13] Carolyn Lamb, G. Brown, Daniel, and L. Clarke, Charles. A taxonomy of generative poetry techniques. Journal of Mathematics and the Arts, 11(3):159–179, 2017.

[14] Alexey Tikhonov and P. Yamshchikov, Ivan. What is wrong with style transfer for texts? In arXiv preprint, 2018.

[15] Tianxiao Shen, Tao Lei, Regina Barzilay, and Tommi Jaakkola. Style transfer from non-parallel text by cross-alignment. 31st Conference on Neural Information Processing Systems, pages 6833–6844, 2017.

[16] Juncen Li, Robin Jia, He He, and Percy Liang. Delete, retrieve, generate: A simple approach to sentiment and style transfer. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, volume 1, pages 1865–1874, 2018.

[17] Rico Sennrich, Barry Haddow, and Alexandra Birch. Controlling politeness in neural machine translation via side constraints. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 35–40, 2016.

[18] James M. Hughes, Nicholas J. Foti, David C. Krakauer, and Daniel N. Rockmore. Quantitative patterns of stylistic influence in the evolution of literature. Proceedings of the National Academy of Sciences, 109(20):7682–7686, 2012.

[19] Harsh Jhamtani, Varun Gangal, Eduard Hovy, and Eric Nyberg. Shakespearizing modern language using copy-enriched sequence-to-sequence models. In Proceedings of the Workshop on Stylistic Variation, pages 10 – 19, 2017.

[20] Keith Carlson, Allen Riddell, and Daniel Rockmore. Zero-shot style transfer in text using recurrent neural networks. In arXiv preprint, 2017.

[21] Zhenxin Fu, Xiaoye Tan, Nanyun Peng, Dongyan Zhao, and Rui Yan. Style transfer in text: Exploration and evaluation. In arXiv preprint, 2017.

[22] Kelvin Guu, Tatsunori B. Hashimoto, Yonatan Oren, and Percy Liang. Generating sentences by editing prototypes. In arXiv preprint:1709.08878, 2017.

[23] Jessica Ficler and Yoav Goldberg. Controlling linguistic style aspects in neural language generation. In Proceedings of the Workshop on Stylistic Variation, pages 94 – 104, 2017.

[24] Remi Lebret, David Grangier, and Michael Auli. Neural text generation from structured data with application to the biography domain. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 1203–1213, 2016.

[25] Alec Radford, Rafal Jozefowicz, and Ilya Sutskever. Learning to generate reviews and discovering sentiment. In arXiv preprint, 2017.

[26] Jian Tang, Yifan Yang, Sam Carton, Ming Zhang, and Qiaozhu Mei. Context-aware natural language generation with recurrent neural networks. In arXiv preprint, 2016.

[27] Zhiting Hu, Zichao Yang, Xiaodan Liang, Ruslan Salakhutdinov, and Eric P. Xing. Toward controlled generation of text. In International Conference on Machine Learning, pages 1587–1596, 2017.

[28] Alexey Tikhonov, Viacheslav Shibaev, Aleksander Nagaev, Aigul Nugmanova, and Ivan P. Yamshchikov. Style transfer for texts: Retrain, report errors, compare with rewrites. In arXiv preprint:1908.06809, 2019.

[29] Ivan P Yamshchikov, Viacheslav Shibaev, Aleksander Nagaev, Jürgen Jost, and Alexey Tikhonov. Decomposing textual information for style transfer. arXiv preprint arXiv:1909.12928, 2019.

[30] Jiwei Li, Michel Galley, Chris Brockett, Georgios P. Spithourakis, Jianfeng Gao, and William B. Dolan. A persona-based neural conversation model. CoRR, abs/1603.06155, 2016.

[31] Jing He, Ming Zhou, and Long Jiang. Generating chinese classical poems with statistical machine translation models. In AAAI, 2012.

[32] Xiaoyuan Yi, Ruoyu Li, and Maosong Sun. Generating chinese classical poems with rnn encoder-decoder. In Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data, pages 211–223, 2017.

[33] Rui Yan. i, poet: Automatic poetry composition through recurrent neural networks with iterative polishing schema. In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16), pages 2238–2244, 2016.

[34] Rui Yan, Cheng-Te Li, Xiaohua Hu, and Ming Zhang. Chinese couplet generation with neural network structures. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pages 2347 – 2357, 2016.

[35] Jiyuan Zhang, Yang Feng, Dong Wang, Yang Wang, Andrew Abel, Shiyue Zhang, and Andi Zhang. Flexible and creative chinese poetry generation using neural memory. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, volume 1, pages 1364–1373, 2017.

[36] Marjan Ghazvininejad, Xing Shi, Yejin Choi, and Kevin Knight. Generating topical poetry. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 1183–1191. Association for Computational Linguistics, 2016.

[37] Peter Potash, Alexey Romanov, and Anna Rumshisky. Ghostwriter: Using an lstm for automatic rap lyric generation. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 1919–1924. Association for Computational Linguistics, 2015.

[38] Yoshua Bengio, Réjean Ducharme, Pascal Vincent, and Christian Jauvin. A neural probabilistic language model. Journal of machine learning research, 3(Feb):1137–1155, 2003.

[39] Frederic Morin and Yoshua Bengio. Hierarchical probabilistic neural network language model. Aistats, pages 246–252, 2005.

[40] Andriy Mnih and Geoffrey Hinton. A scalable hierarchical distributed language model. In Advances in neural information processing systems, pages 1081–1088, 2009.

[41] Yangyang Shi. Language Models with Meta-information. PhD thesis, Technical University of Delft, 2014.

[42] Li Fei-Fei, Rob Fergus, and Pietro Perona. One-shot learning of object categories. In IEEE transactions on pattern analysis and machine intelligence, pages 594 – 611, 2006.

[43] Sebastian Thrun and Lorien Pratt. Learning to learn. Springer Science & Business Media, 2012.

[44] Douglas Bagnall. Author identification using multi-headed recurrent neural networks. In arXiv preprint, 2015.

[45] Senko K. Maynard. The centrality of thematic relations in japanese text. Functions of language, 1(2):229–260, 1994.

[46] Aleksei Kruchenykh. Phonetics of theater. M.:41, Moscow, 1923.

[47] Victor Shklovsky. Poetics: on the theory of poetic language. 18 State typography, Petrograd, 1919.

[48] E Blasi, Damián, Søren Wichmann, Harald Hammarström, F. Stadler, Peter, and H. Christiansen, Morten. Sound – meaning association biases evidenced across thousands of languages. Proceedings of the National Academy of Sciences, 113(39):10818 – 10823, 2016.

[49] Ivan Yamshchikov, Viacheslav Shibaev, and Alexey Tikhonov. Dyr bul shchyl. proxying sound symbolism with word embeddings. In Proceedings of the 3rd Workshop on Evaluating Vector Space Representations for NLP, pages 90–94, 2019.

[50] Alexey Tikhonov and P. Yamshchikov, Ivan. Music generation with variational recurrent autoencoder supported by history. In arXiv preprint, 2017.

[51] Rafal Jozefowicz, Oriol Vinyals, Mike Schuster, Noam Shazeer, and Yonghui Wu. Exploring the limits of language modeling. In arXiv preprint, 2016.

[52] Ziang Xie. Neural text generation: A practical guide. In arXiv preprint, 2017.

[53] Bo-June Hsu and James Glass. Iterative language model estimation: efficient data structure & algorithms. In Ninth Annual Conference of the International Speech Communication Association, 2008.

[54] L.A. Gatys, A. S. Ecker, and M. Bethge. Image style transfer using convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2414 – 2423, 2016.

[55] B. Joshi, K. Stewart, and D. Shapiro. Bringing impressionism to life with neural style transfer in come swim. In arXiv preprint:1701.04928, 2018.

A Examples of output

Table 8 lists some illustrative mistakes of the model both for English and for the Russian language. Reading the raw output, we could see several types of recurring characteristic errors that are typical for LSTM-based text generation. They can be broadly classified into several different types:

- If the target author is underrepresented in the training dataset, the model tends to make more mistakes, mostly, syntactic ones;
- Since generation is done in a word-by-word manner, the model can deviate when sampling a low-frequency word;
- Pronouns tend to cluster together, possibly due to the problem of anaphoras in the training dataset;
- The line can end abruptly; this problem also seems to occur more frequently for the authors that are underrepresented in the training dataset.

Table 9 lists some subjectively cherry-picked, especially successful examples of the system outputs both for English and for the Russian language. Since the text is generated line by line, and verses are obtained through random rhyme or rhythm filtering, several types of serendipitous events occur. They can be broadly classified into four different types:

- Wording of the verse that fits into the style of the target author;
- Pseudo-plot that is perceived by the reader due to a coincidental cross-reference between two lines;
- Pseudo-metaphor that is perceived by the reader due to a coincidental cross-reference between two lines;
- Sentiment and emotional ambiance that corresponds to the subjective perception of the target author.

B Cross-entropy for texts in Russian

The cross-entropy between generated texts and samples of human-written texts was calculated as described in Section 7.1. The results are shown in Table 10

One can see that the model achieves the lowest level of cross-entropy when tested on the texts of the target author for all four authors.

C Survey design

The surveys were designed identically for English and Russian languages. We have recruited the respondents via social media; the only prerequisite was fluency in the target language. Respondents were asked to determine authorship for 16 different 4-line verses. The verses for human-written text were chosen randomly out of the data for the given author. The generated verses were obtained through line-by-line automated rhyme and rhythm heuristic filtering. Since LSTMs are not perfect in text generation and tend to have clear problems illustrated in Table 8, we additionally filtered generative texts leaving the verses that do not contain obvious mistakes described above. Each of the 16 questions consisted of a text (in lower case with stripped-off punctuation) and multiple-choice options listing five authors, namely, four human authors and an artificial neural network. Respondents were informed that they are to distinguish human- and machine-written texts. The correct answers were not shown to the respondents. Table 7 shows the results of the survey for English texts and Table 11 for Russian ones. Higher values in every row correspond to the options that were more popular among the respondents when they were presented with the text written by the author listed in the first column of the table.

Table 8: Examples of several recurring types of mistakes that occur within generated lyrics.

Problem	English	Russain
broken syntax	$Neuro-MUSE$: every step inside on our faces that i would stop my self going crazy	$Neuro-Zemfira$: ты слышишь я слышу шаги мои в душу дрожи
'rare' word brakes the line	$Neuro-Shakespeare$: o ho de profundis she says i am on her	$Neuro-Letov$: иду гляжу в окно гляжу мне вслед на небо
pronouns 'entangle'	$Neuro-Shakespeare$: thou here shalt be and thine who will have to my grave	$Neuro-Zemfira$: тебе ли ты ль за
sentences don't end	$Neuro-Muse$: at night i lay waiting for a	$Neuro-Brodsky$: двух четырех десять лет за углом
nonsense	$Neuro-Shakespeare$: do many a fair honour best of make or lose	$Neuro-Lenov$: о о о о и о

Table 9: Cherry-picked examples of generated lyrics after either rhyme or rhythm filtering illustrating typical serendipities.

Serendipity	English	Russain
peculiar wording	$Neuro-Shakespeare$: a sense i may not comprehend of whom i had not to defend	$Neuro-Pushkin$: во славу вакха или тьмы мы гордо пировали
apophenic plot	$Neuro-Marley$: oh lord i know how long i d burn take it and push it it s your turn	$Neuro-Esenin$: ты под солнцем стоишь и в порфире как в шелку беззаботно горишь
apophenic metaphor	$Neuro-Carroll$: your laugh is bright with eyes that gleam that might have seen a sudden dream	$Neuro-Zemfira$: ветер в голове с красной тенью шепчется
peculiar sentiment	$Neuro-Muse$: i am the man of this universe i remember i still am a curse	$Neuro-Letov$: только в ушах отражается даль только белая смерть превращается в ад

Table 10: Sample cross entropy between generated texts $\{T_i^G|A_i\}$ and actual texts for different authors. The two smallest values in each row are marked with * and ** and a bold typeface. The sample cross entropy between random samples from the texts of the target author and randomly generated sequences of words (uniform and weighted respectively) as well as other samples written by the same author (denoted as SELF) are shown for reference.

Model $G(A_i)$/ author	Pushkin	Esenin	Brodsky	Letov	Zemfira
Neuro-Pushkin	**17.9***	**21.8****	23.4	27.0	30.8
Neuro-Esenin	**20.4****	**18.8***	21.0	22.7	26.0
Neuro-Brodsky	23.5	**21.1****	**17.2***	20.9	23.8
Neuro-Letov	22.2	**20.0****	20.8	**19.6***	23.6
Neuro-Zemfira	19.5	**17.1****	18.1	18.2	**16.6***
Uniform Random	103.0	103.1	103.0	103.0	103.8
Weighted Random	40.8	40.2	40.2	42.6	45.6
SELF	35.0	33.7	38.0	28.3	12.0

Table 11: Results of a survey with 178 respondents. Shares of each out of 5 different answers given by people when reading an exempt of a poetic text by the author stated in the first column. The two biggest values in each row are marked with * and ** and a bold typeface.

Model $G(A_i)$ or author	Pushkin	Esenin	Letov	Zemfira	LSTM
Neuro-Pushkin	**0.31****	0.22	0.02	0.0	**0.44***
Pushkin	**0.62***	0.11	0.03	0.01	**0.23****
Neuro-Esenin	0.02	**0.61***	0.08	0.0	**0.29****
Esenin	0.06	**0.56***	0.07	0.02	**0.29****
Neuro-Letov	0.0	0.02	**0.40****	0.08	**0.51***
Letov	0.0	0.01	**0.61***	0.02	**0.35****
Neuro-Zemfira	0.0	0.06	0.13	**0.4****	**0.41***
Zemfira	0.0	0.02	0.08	**0.58***	**0.31****

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 1, 119-127 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Performance Portability and Unified Profiling for Finite Element Methods on Parallel Systems

Vladyslav Kucher[*,1], Jens Hunloh[2], Sergei Gorlatch[2]

[1]*National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Prosp. Peremohy 37, Kyiv, 03056, Ukraine*

[2]*University of Muenster, Einsteinstr. 62, Muenster, Germany*

ARTICLE INFO

Article history:
Received: 11 November, 2019
Accepted: 20 December, 2019
Online: 24 January, 2020

Keywords:
C++ Compilers
GPU programming
High-performance computing
Performance portability
Finite element methods
Parallel computing
Portable profiling
Unified parallel programming

ABSTRACT

The currently available variety of modern, highly-parallel universal processors includes multi-core CPU and many-core GPU (Graphics Processing Units) from different vendors. Systems composed of such processors enable high-performance execution of demanding applications like numerical Finite Element Methods. However, today's application programming for parallel systems lacks performance portability: the same program code cannot achieve stable high performance on different parallel architectures. One of the main reasons for this is that parallel programs are developed by utilizing system-specific profiling interfaces of the corresponding hardware vendors. We describe a novel, portable profiling interface: its design, implementation, and evaluation within the popular framework DUNE for solving differential equations using finite element methods. Our profiler is built on top of the PACXX framework for parallel programming in C++, and it supports portable parallel programming using a single profiling tool on various target hardware platforms.

1 Introduction

Modern universal processors are becoming highly parallel: multi-core CPU and many-core GPU (Graphics Processing Units) are produced by different vendors and exemplify a variety of architecture and hardware solutions. Such processors enable building high-performance systems for computation-intensive applications like numerical finite element methods.

The current programming approaches for parallel computing systems include CUDA [1] that is restricted to GPU produced by NVIDIA, as well as more universal programming models - OpenCL [2], SYCL [3], and PACXX [4] - that follow the idea of unified programming: the programmer can target different hardware without changing the source code, thus reducing the development overhead. However, existing profiling tools are still restricted to the corresponding vendor; therefore, the application programmer usually has to use several different tools to achieve the ultimate portability.

There have been several research efforts to make profiling tools more portable and flexible. CUDA Advisor [5] is a profiling interface that collects data about the performance of both CPU and GPU parts of program code by using the LLVM infrastructure [6]. Experimental approach [7] uses CUDA's TAU tool for profiling GPU applications by producing a detailed information on communications between CPU and GPU, without modifying the source code. The SASSI instrumentation tool (NVIDIA assembly code "SASS Instrumentor") [8] relies on customizable metrics: it enables to inserting user-defined instrumentation code (e.g., debug hooks and customized counters) to collect detailed profiling data that are usually not available. The generic tool integrated in the Score-P infrastructure [9] provides an interface for evaluating the performance of OpenCL code on different parallel architectures.

This paper aims at designing a single, cross-platform profiler that significantly improves the portability of program development. We design our profiler by extending the unified programming framework *PACXX (Programming Accelerators in C++)* [4] with a generic profiling interface: the interface follows the unified programming model of PACXX. Our profiler seamlessly extends the PACXX programming system and enables collecting profiling information at different stages of the development process, for different kinds of the target hardware, thus reducing the development overhead.

[*]Vladyslav Kucher, Prosp. Peremohy 37, Kyiv, 03056, Ukraine, commastergm@gmail.com

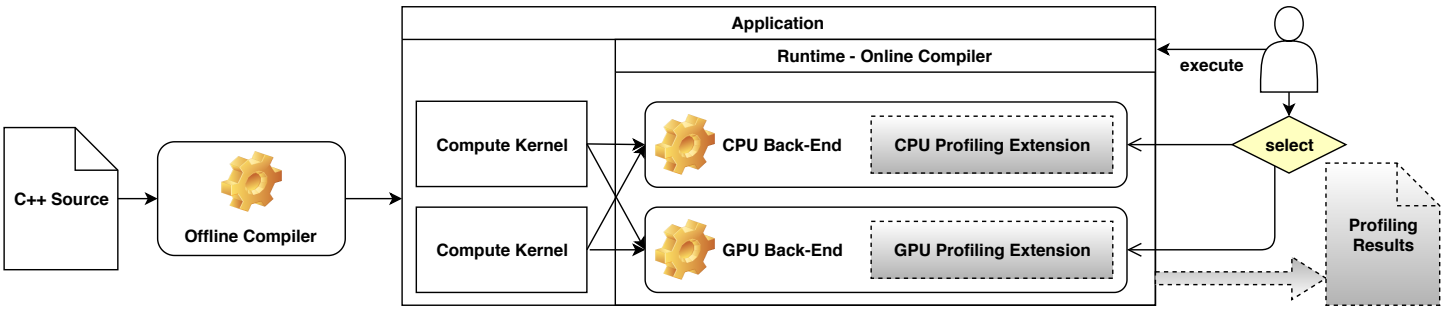


Figure 1: The PACXX framework, extended with profiling (shaded parts)

The unified parallel programming approach of PACXX enables programming CPU/GPU systems in the newest standard C++14/17. PACXX significantly simplifies the development of portable parallel C++ programs by slightly modifying the source code staying within C++. The PACXX advantages were demonstrated for application areas like linear algebra, stencils, and N-Body simulations [10].

We choose *Partial Differential Equations (PDE)* solving as an application field to illustrate and evaluate our approach to unified profiling, because PDE are broadly used in different areas of engineering and science, including simulation [11], finance [12], numerical modeling [13], and biology [14]. Most of practically important PDE are solved using finite-element (FE) methods. Unfortunately, these methods' implementation is often tied to a particular data structure representing the corresponding computation grid. C++ is often used for implementing finite-element methods, because of its high performance and modularity. The *Finite Element Computational Software (FEniCS)* [15] is a C++-based framework for FE variational problems, relying on automatic code generation and the custom *FEniCS Form Compiler* [16]: it supports performance portability by employing a domain-specific language based on Python [17], and multi-staging optimization. The Firedrake approach [18] achieves performance portability by using an advanced code generator: problems are specified in a Python-embedded, domain-specific language that is lowered to C and then JIT-compiled, so target code is generated at run time. The object-oriented framework Overture [19] for solving PDE on various platforms contains C++ classes for computations in different geometries. The Feel++ framework [20] uses a C++-based domain-specific language whose implementation is compatible with third-party libraries.

In order to evaluate our approach to unified profiling, we consider the *Distributed and Unified Numerics Environment (DUNE)* [21] as a case study. DUNE offers a unified interface for solving PDE on parallel systems [22]: the interface combines various methods in a modular C++ library. DUNE enables using arbitrarily shaped grids and a large set of particular algorithms. While the design of DUNE is very flexible, its application performance strongly depends on the employed C++ compiler and the target hardware.

The goal of this paper is to extend the flexibility of existing approaches to parallel PDE solving. We achieve performance portability and cross-platform profiling of the finite-element methods for solving PDE, by extending the PACXX framework and integrating it with the existing DUNE framework.

The paper is organized as follows. In Section 2, we describe the C++-based unified programming approach of PACXX and the design of our profiling interface. In Section 3, we show how this interface is used for a portable profiling a simple case on various target architectures. In Sections 4 and 5, we illustrate how PACXX can accelerate grid-based FE algorithms for PDE in the extended DUNE. We demonstrate that the integration of PACXX with DUNE leads to performance portability over different architectures, significantly decreasing the development overhead and improving code maintenance. We experimentally evaluate our unified profiling approach in Section 6, and we summarize our results in Section 7.

2 Programming and profiling

2.1 The PACXX Programming model

We use the PACXX framework [4] that supports a unified, C++-based model of parallel programming. In analogy with the recent OpenCL standard [23], PACXX uses the parallelism expressed using kernels that are executed on devices. The main advantage of PACXX, however, is that an application is expressed as single-source C++ code, whereas OpenCL kernels are implemented in a special kernel language and they need additional host code to be executable. The performance portability in PACXX is ensured by means of several pre-implemented back-ends.

Figure 1 shows that the compilation process of PACXX proceeds in two steps. The first step is the offline compilation that transforms source code into an executable. Kernel's code is precompiled, and the executable is integrated with the particular back-end for the target CPU or GPU. In the second stage, online compilation is invoked when the executable is started. The integration of the executable with back-ends for different architectures allows the user to choose the target hardware for each execution. The efficient execution on the chosen target system is handled by PACXX transparently: all necessary generation steps and optimizations are performed automatically. This brings the portability of PACXX applications on CPUs and GPUs of different vendors.

2.2 Profiling interface: Design and implementation

Figure 1 shows the design of our unified profiling interface integrated into PACXX: shaded are our profiling extensions to the original PACXX back-ends. These extensions comprise hardware-specific code to collect profiling data. For the user, our profiling interface can be viewed as a wrapper that can choose the corresponding profiling extension for the particular target architecture.

The advantage of our design of profiling is that no changes to the offline compiler of PACXX are necessary. The modifications are made only to the PACXX runtime and back-ends: they are extended with vendor-specific profiling tools: the CPU back-end uses the PAPI library [24], and the back-end for NVIDIA GPUs relies on the CUPTI library [25].

2.3 The profiling extension

For different back-ends, the profiling extensions have a unified structure that includes a changed launch procedure for kernels to manage collecting performance data. We have to modify the way how a kernel is launched, because many existing devices cannot record several profiling metrics simultaneously (e.g., NVIDIA GPUs [25]). For some metrics, multiple reproducible runs of the same kernel may be necessary to record that metric. It becomes mandatory to rebuild the device memory state automatically after each kernel execution, such that the executions are measured independently. This reconstruction of the memory state enables multiple profiling runs for the same program, so that accurate performance data are produced. Traditionally, this is done using a shadow copy mechanism that stores and restores the device memory state transparently at each kernel execution. This solution effectively decouples the profiling from undesirable side effects: eventually this ensures exact performance data and deterministic behavior when measuring multiple kernel launches.

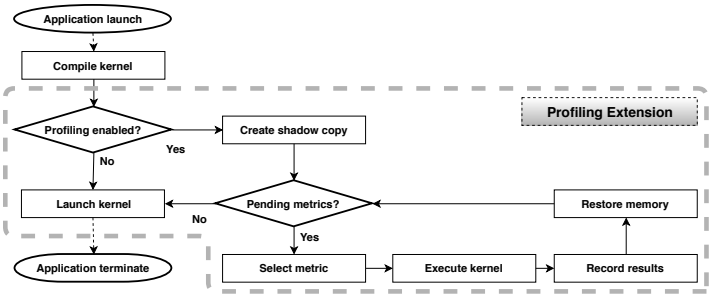


Figure 2: Kernel launch, modified for profiling

Figure 2 shows our modification (in dashed box) to the launch procedure within a PACXX back-end. The new launch procedure reflects the control flow of our profiling extension: the processing steps rely on target-specific code of the concrete back-end. When profiling is enabled, a kernel launch triggers the profiling extension to store a shadow copy of the utilized part of device memory. Then the kernel is executed repeatedly, such that this shadow copy is used for every recorded performance metric. Eventually, the application proceeds using the original kernel launch procedure.

3 The profiler usage

We demonstrate the usage of and evaluate our unified profiler on two application examples - the first is traditional multiplication of matrices, and the second is a more elaborated PDE solver.

3.1 Example: Matrix multiplication

Figure 3 illustrates how matrix multiplication is expressed in PACXX [4]: this program is created from the C++ code by transforming nested sequential loops into parallel calls of a kernel. Arrays *a* and *b* stand for the input matrices, and array *c* stores the resulting matrix. The *matrixMultKernel* kernel (lines 15-23) is written as a C++ lambda expression to be computed in parallel for each element of *c*. The kernel run requires an upload of the input data (lines 3-13) and the result fetching (line 26) after the computation, which is analogous to CUDA and OpenCL kernels.

```
1  auto& device = Executor::get(0);
2
3  auto& da = device.allocate<double>(matrix_size);
4  auto& db = device.allocate<double>(matrix_size);
5  auto& dc = device.allocate<double>(matrix_size);
6
7  da.upload(a, matrix_size);
8  db.upload(b, matrix_size);
9  dc.upload(c, matrix_size);
10
11 auto pa = da.get();
12 auto pb = db.get();
13 auto pc = dc.get();
14
15 auto matrixMultKernel = [=](auto &config)
16 {
17    auto col = config.get_global(0);
18    auto row    = config.get_global(1);
19    double val  = 0;
20    for (unsigned i = 0; i < width; ++i)
21       val += pa[row * width + i] *
                pb[i * width + col];
22    pc[row * width + col] = val;
23 };
24 exec.launch(matrixMultKernel,
     {{width/threads , width} , {threads , 1} , 0});
25
26 dc.download(c, matrix_size);
```

Figure 3: PACXX code example: matrix multiplication

Figure 3 shows that the code employs parallelism by dividing calculations in blocks and threads. Function *launch* (line 24) invokes the kernel and it partitions work in dimensions (up to three), thus specifying the structure and degree of parallelism. In the range, the first component declares how many blocks are started, while the *threads* variable in the range declares the threads number in a block. At run time, each block is assigned to a processor that runs all threads of the block in the SIMT (Single Instruction, Multiple Threads) manner, which is a combination of SIMD and multi-threading paradigms. Function *get_global* retrieves the thread id in each dimension of the range. The first dimension of the range in our example corresponds to the row, and the second to the column in the matrix.

After the invocation of PACXX offline compiler, the executable is automatically combined with our profiler. Therefore, PACXX serves effectively as a drop-in for the standard LLVM (Low-Level Virtual Machine) [6] compilation toolchain.

3.2 Workflow of profiling

In Figure 4, the profiling of our example application is invoked with the runtime environment variable PACXX_PROF_ENABLE enabled, which leads to profiling each kernel invocation. At each kernel run, the runtime system executes the online compilation transparently for the user and it enables the profiler, such that the profiling information is collected from the target device. Variable PACXX_DEFAULT_RT specifies the target architecture (CPU or GPU). Profiling data are written in an output file that can be either explicitly specified as in Figure 4 or sent to the standard output.

```
> PACXX_DEFAULT_RT=GPU
  PACXX_PROF_ENABLE=1
  PACXX_PROF_IN=custom_profiling_configuration
  PACXX_PROF_OUT=pacxx.json
  ./matrixMult
```

Figure 4: Invoking a PACXX program with profiling

Our profiler always reports one metric - runtime duration - by default. To ask for additional profiling data, the user can specify a specific configuration. The profiler is able to profile all metrics that are supported by the system vendor, e.g., memory throughput and usage, instructions per cycle, etc. All 174 possible metrics for NVIDIA GPUs are listed in [25].

3.3 Visualization of profiling

Figure 5 illustrates the profiling data collected for the matrix multiply example prepared as shown in Figure 4. The profiler produces a JSON-format file that comprises all specified metrics for each started kernel instance. For multiple kernel launches, our profiler records every launch independently, and the output data are prefixed by the kernel name. Thus, the user obtains profiling information automatically: the user does not have to explicitly select the profiling tool for a kernel and the architecture on which it runs: the appropriate profiling functionality is ensured by the PACXX runtime for the target architecture.

```
"matrixMultKernel": [
  { "Metrics": {
      "cf_executed": "68157540",
      "kernelDuration": "517349264ns" } } ]
```

Figure 5: Results of profiling in JSON format

The use of the JSON format enables efficient visualization of the profiling results using Gnuplot or other tools, such that the code behavior on different target platforms can be conveniently compared.

Figure 6 compares the execution time of matrix multiplication on a CPU (Intel Xeon E5-1620 v2) vs. a GPU (NVIDIA Tesla K20c). The input comprises two 4096x4096 square matrices generated randomly. The plot demonstrates how the execution time is dependent of the parallelism amount specified by parameter *threads* in the code in Figure 3.

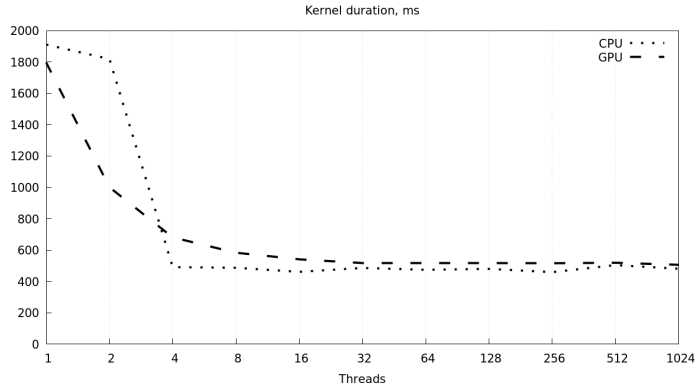


Figure 6: Execution time of matrix multiply on CPU vs. GPU

Figure 7 compares the number of branch statements on CPU vs. GPU, depending on the number of threads.

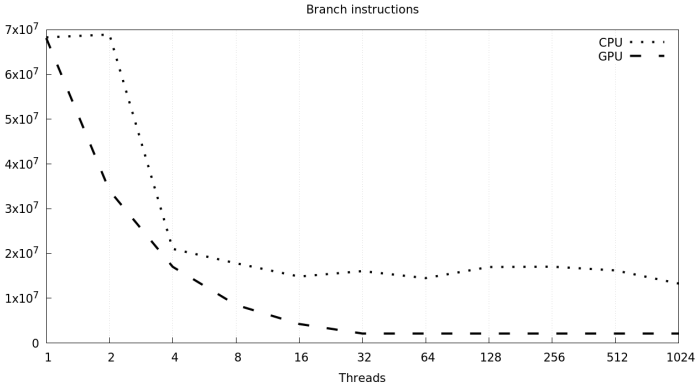


Figure 7: Number of branch instructions on CPU vs. GPU

Summarizing, our profiler enables both specifying metrics and representing profiling results in a uniform manner, such that different target architectures can be compared with each other conveniently.

4 Integrating DUNE with PACXX

In this section, we show the use of our approach to profiling on the case study of the DUNE framework. We explain its efficient integration with the PACXX unified programming framework, by taking into account the design of DUNE's grid interface.

Our idea of integration is to employ PACXX as a code generator for the DUNE interface which is designed by using abstract interfaces and template meta-programming to separate data structures and algorithms. By reusing existing modules, DUNE offers a rich collection of functionalities.

In our integration with PACXX, we use the fact that DUNE contains C++ modules that provide abstract interfaces with several implementations. The use of generic programming [26] and C++ template meta-programming [27] enables optimizations to be applied at compile time: therefore, DUNE is flexible, without introducing overhead at runtime.

Figure 8 illustrates that our integration comprises two levels. At the top (abstraction) level, a particular problem is formulated

using extensions of a domain-specific language analogous to [28], and is then implemented as a specific DUNE kernel. At the bottom (hardware optimization) level, code is generated for a specific target architecture. This two-level design of the integration toolchain enables suitable optimizations at every processing level.

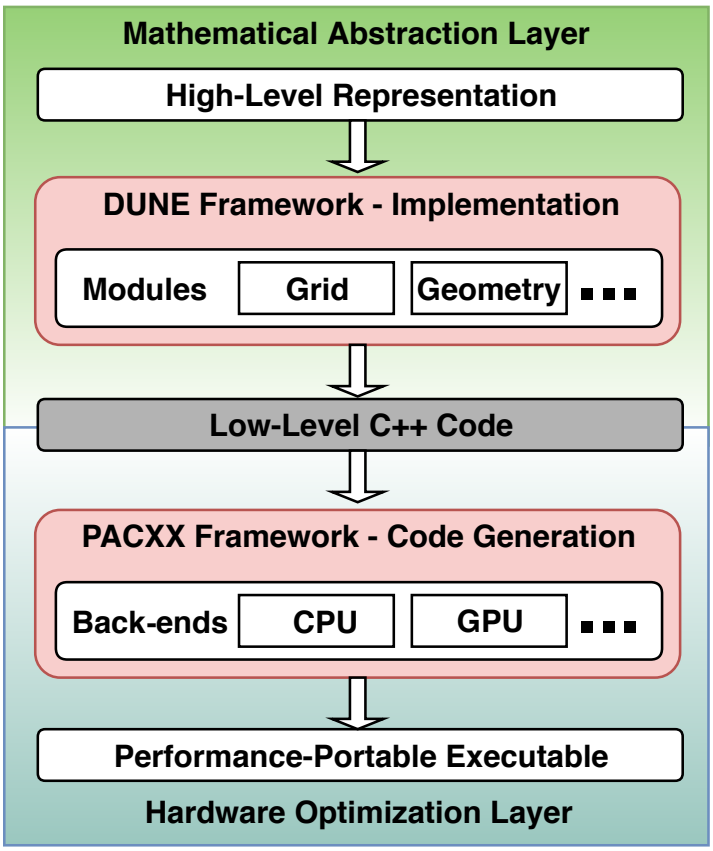


Figure 8: Integration of PACXX with DUNE

The integrated toolchain is now as follows. The PDE problem is stated by the domain expert as an expression of a domain-specific language, which is then implemented by a specific grid-based kernel of DUNE: the framework yields C++ code that contains the resulting kernel with template arguments and runtime values. This code is eventually compiled by PACXX, as we explain below.

The PACXX framework employs the popular LLVM compiler infrastructure and its Intermediate Representation (IR) [6] to obtain portable executables with hardware-specific optimizations. Kernels are compiled offline to IR and then the online compiler generates code and optimizes it for the target architecture [10] transparently for the user. This toolchain enables portable PACXX applications over architectures of modern parallel processors. PACXX also supports C++ template meta-programming which significantly reduces the runtime overhead, and our unified profiling allows profiling of the DUNE without further modifications to the code.

The main advantage of using PACXX vs. existing profilers is the profiling of applications over various platforms, without having to use different tools. Our profiler improves the flexibility of programming process: the program can be configured regarding requested profiling data on every supported target architecture. In addition, due to the use of PACXX for generating code in DUNE we can apply optimizations based on control-flow linearization [29]. Since PACXX is fully compatible with the modern C++14/17 standards, the program developers can use existing C++ code on various target architectures, which significantly reduces the development effort.

5 Compiling kernels: examples

This section describes the use of our integrated DUNE-PACXX framework for generating code for two examples of DUNE kernels: 1) Poisson equation solving, and 2) deformation of material calculation based on the linear elasticity theory. These two examples cover two scenarios: Poisson calculates with low load on a simple 2D grid, and linear elasticity is based on a 3D grid with a heavy load. Example applications of linear elasticity are in computer-aided design, and Poisson is applied in physics to calculate fields of potentials.

5.1 PACXX parallelization

We consider example kernels that are simple sequential C++ codes; they are parallelized according to the PACXX programming approach. The slight one-time adaptation effort of the kernels leads to the performance portability of the resulting code. Thereby the programmer does not have to re-design the program code for each new parallel architecture.

Our two kernel examples work on a grid refined on two levels: the computation in a uniform grid of *macro-elements* uses a coarse-grained refinement, and each macro-element is further refined to a grid of *micro-elements*. We parallelize kernels by two ways: either different macro-elements are processed in parallel, or several micro-elements in a macro-element are processed in parallel.

We compare the micro- vs. macro-parallelization approaches regarding their performance against an original sequential version.

5.2 The Poisson kernel

```
1  // iterating over micro-elements
2  for (int el_y=0; el_y<k; ++el_y)
3   for (int el_x=0; el_x<k; ++el_x)
4   {
5    // (1) calculating
6    RF grad_u[2] = {};
7    for (int dim=0; dim<=1; ++dim)
8    {
9     for (int index_y=0; index_y<2; ++index_y)
10     for (int index_x=0; index_x<2; ++index_x)
11      grad_u[dim] += grad[index_x+index_y*2][0][dim]
        * x(lfsv, (el_y+index_y)*(k+1) + el_x + index_x);
12   }
13   // (2) accumulating
14   for (int index_y=0; index_y<2; ++index_y)
15    for (int index_x=0; index_x<2; ++index_x)
16     r.accum(lfsv,
       (el_y+index_y)*(k+1) + el_x + index_x,
       factor*(grad[index_x+index_y*2][0][0]*grad_u[0]
              +grad[index_x+index_y*2][0][1]*grad_u[1]));
17 }
```

Figure 9: Poisson kernel: sequential C++ code

Figure 9 illustrates the C++ code for the Poisson kernel that computes a single macro-element in DUNE. Two nested *for* loops in

lines 2-3 iterate on the micro-grid. For every element, the computation comprising the loop body is executed. Local basis functions $lfsv$ and x in line 11 compute a gradient, and the consequent accumulation phase (line 16) uses the calculated gradient for updating the computational grid.

Figure 10 shows how the memory is accessed by the kernel in Figure 5.

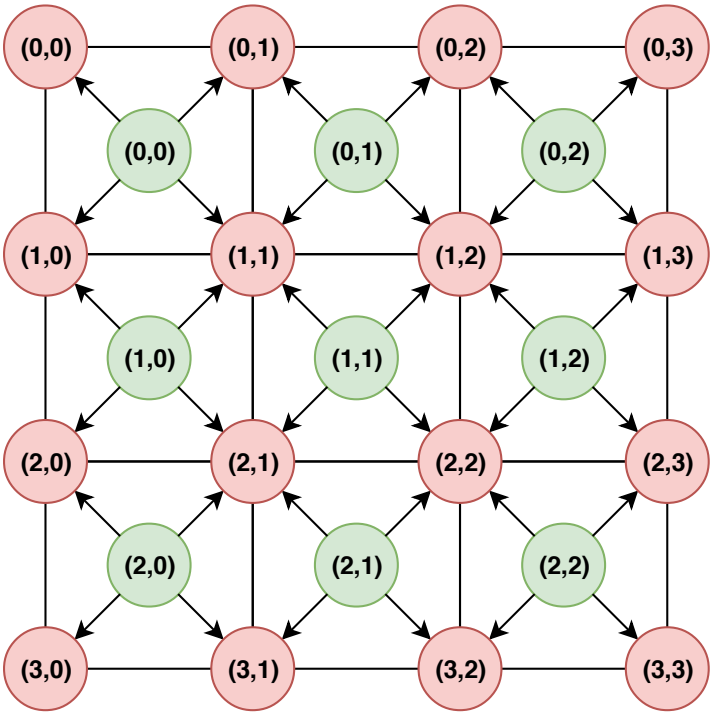


Figure 10: Poisson: micro-grid and data dependencies

For correct parallelization, we must avoid race conditions. Therefore, we synchronize accesses to memory by distinguishing between two kernels - accumulation and computation. We parallelize the calculation kernel on the micro-grid elements, but temporarily store the results of the calculation kernel, rather than performing write operations immediately.

Afterwards, the corresponding vertices are processed in parallel by the accumulation kernel: it reads the previously stored intermediate results and updates the adjacent vertices by simultaneous write operations. This described implementation guarantees the absence of conflicts in the patterns of memory access at the level of micro-grid.

At a macro-layer, we parallelize computations by executing the main computation for each instance of the kernel, such that process in parallel all macro-elements. For this approach, DUNE's grid interface must be slightly modified, but the advantage is that the patterns of memory access are not so complicated.

The parallelization at macro-layer implies that workload is distributed in a coarse-grained manner; this is advantageous on large-scale since it greatly reduces the parallelization overhead. However, it brings a drawback: global memory must be allocated by all macro-elements simultaneously, so the memory requirement increases. Vice versa, the fine-grained distribution of workload for micro-layer parallelism allocates little memory for a macro-element at a time. Therefore, large micro-grids with enough compute load in a single macro-element especially benefit from the micro-layer parallelization.

5.3 Kernel for linear elasticity

Figure 11 shows our second example - the linear elasticity kernel – that consists of three nested *for* loops. and is thus significantly more complex than Poisson. On the one hand, this complicates the parallelization (while the scheme of parallelization remains similar to Poisson), but on the other hand, this offers more potential of parallelism.

```
1  // iterating over micro-elements
2  for (int el_z = 0; el_z < k; ++el_z)
3   for (int el_y = 0; el_y < k; ++el_y)
4    for (int el_x = 0; el_x < k; ++el_x)
5    {// (1) calculating
6     for (int dim = 0; dim <= 2; ++dim)
7      for (int index_z = 0; index_z <= 1; ++index_z)
8       for (int index_y = 0; index_y <= 1; ++index_y)
9        for (int index_x = 0; index_x <= 1; ++index_x)
10       {// computating the element values ...}
11    // (2) accumulating
12    for (int index_z = 0; index_z <= 1; ++index_z)
13     for (int index_y = 0; index_y <= 1; ++index_y)
14      for (int index_x = 0; index_x <= 1; ++index_x)
15       {// accumulation of node values...}
```

Figure 11: Linear elasticity: sequential C++ code

6 Experimental evaluation

We conduct our measurements using two parallel processors: an Intel CPU Xeon E5-1620v2 with 4 cores of frequency 3.70GHz (8 logical cores), and an NVIDIA GPU Tesla K20c with 2496 compute units. Our CPU contains AVX with vector registers, so we evaluate both scalar and auto-vectorized versions of our kernels. The number of (sub)elements (x-axis) corresponds to the subdivision degree of the grid in each dimension: e.g., 5 sub-elements imply a 5x5x5 grid in a 3D program for linear elasticity.

Each measurement performs one iteration of the linear elasticity program which corresponds to 16 executions of the kernel based on the amount of quadrature points. The following figures compare the three aforementioned platforms (vectorized CPU, scalar CPU and GPU) based on various metrics.

Figure 12 and Figure 13 show the memory operations (lower is better). We observe that read operations can be coalesced quite well, while write operations less so. The GPU with its 32 memory controllers sets the target for other hardware platforms which they struggle to reach.

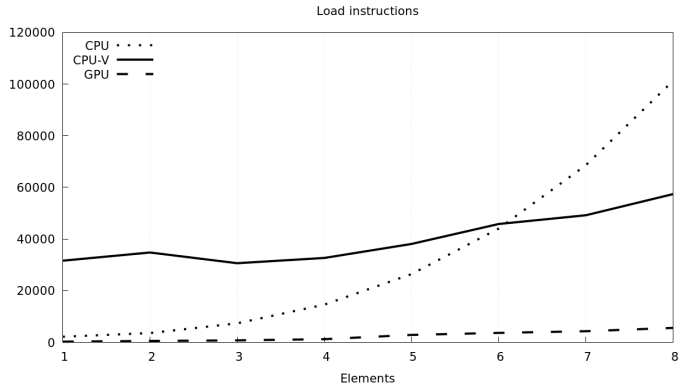


Figure 12: Number of executed memory reads across used platforms

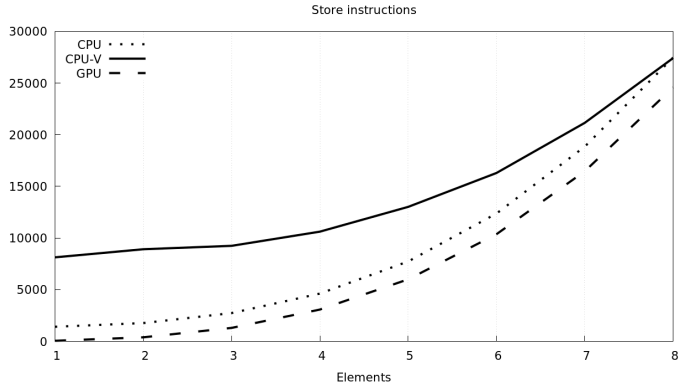


Figure 13: Number of executed memory writes across used platforms

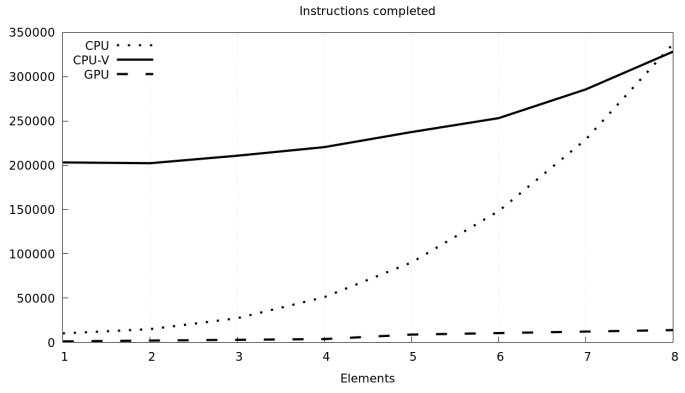


Figure 14: Total completed instructions across used platforms

On the code side, Figure 14 shows the number of executed instructions. We see that CPU vectorization suffers from the low data sizes generating too many scatter-gather instructions (Figure 15 further supports this claim, showing an order of magnitude more processed control-flow instructions in the vectorized CPU version of the kernel). The GPU can distribute the load across its numerous 2496 computing units and keep the number of instructions low.

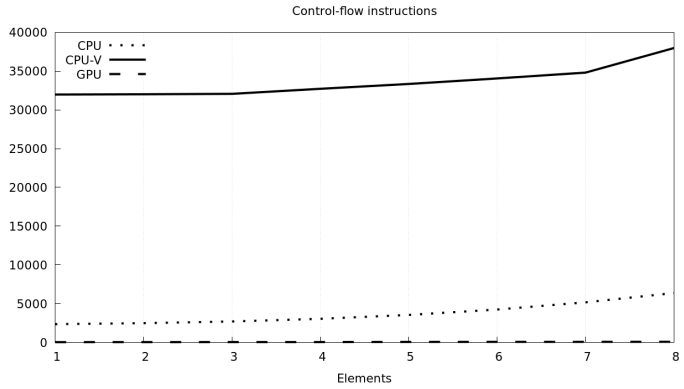


Figure 15: Number of executed control-flow instructions across used platforms

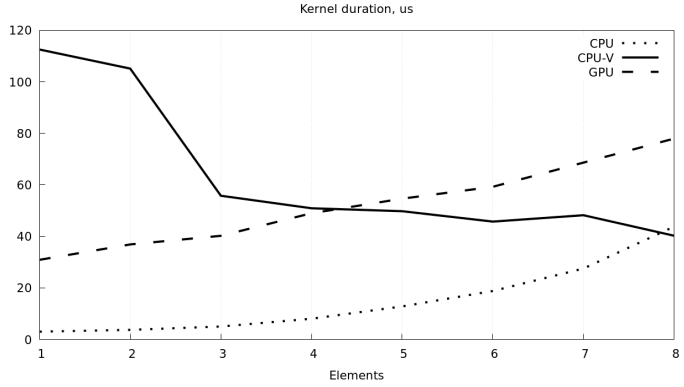


Figure 16: Linear elasticity kernel duration across used platforms

Figure 16 compares the run time of kernel execution. Vectorized CPU version of the kernel does not show good performance with low data amounts, but it improves over the scalar CPU version by element count of 8, matching with the expectation from Figure 14. Surprisingly, the GPU version lags behind the scalar CPU version, so we look into the reasons of low performance on the GPU side.

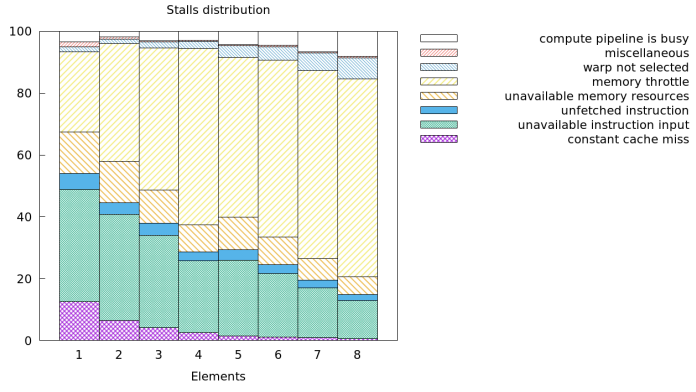


Figure 17: Execution stalls on the GPU causing low performance

Figure 17 shows that the major slowdown reason is memory throttling caused by the hardware limitation to issue memory instructions every 4 cycles. Therefore, memory requests have high divergence and, thus, cannot be fulfilled in a timely manner.

7 Conclusion

We design and implement a novel cross-platform approach to programming and profiling parallel applications, which significantly improves the portability of program development. We extend the unified programming framework PACXX with a generic profiling interface that enables collecting profiling information at different stages of the development process, for different target hardware. Our approach liberates the program developer from having to use several proprietary tools when profiling the same application on different hardware. Therefore, the software development overhead is significantly reduced.

We choose Partial Differential Equations (PDE) solving to illustrate and evaluate our approach of unified profiling, because PDE are broadly used in different areas of engineering and science. As a case study, we use the popular DUNE framework, and we experimentally confirm that the integration of our extended PACXX with DUNE leads to performance portability over different architectures, significantly decreasing development overhead and improving code maintenance.

Acknowledgement The authors gratefully acknowledge generous support from the German Federal Ministry of Education and Research (BMBF) within the HPC^2SE project.

References

[1] J. Nickolls, I. Buck, M. Garland, K. Skadron, "Scalable Parallel Programming with CUDA," *Queue*, vol. 6, pp. 40–53, Mar. 2008.

[2] J. Kim, S. Seo, J. Lee, J. Nah, G. Jo, J. Lee, "OpenCL As a Unified Programming Model for Heterogeneous CPU/GPU Clusters," *SIGPLAN Not.*, vol. 47, pp. 299–300, Feb. 2012.

[3] R. Keryell, R. Reyes, L. Howes, "Khronos SYCL for OpenCL: A Tutorial," in *Proceedings of the 3rd International Workshop on OpenCL*, IWOCL '15, (New York, NY, USA), pp. 24:1–24:1, ACM, 2015.

[4] M. Haidl, M. Steuwer, T. Humernbrum, and S. Gorlatch, "Multi-stage Programming for GPUs in C++ Using PACXX," in *Proceedings of the 9th Annual Workshop on General Purpose Processing Using Graphics Processing Unit*, GPGPU '16, (New York, NY, USA), pp. 32–41, ACM, 2016.

[5] D. Shen, S. L. Song, A. Li, X. Liu, "CUDAAdvisor: LLVM-based Runtime Profiling for Modern GPUs," in *Proceedings of the 2018 International Symposium on Code Generation and Optimization*, CGO 2018, (New York, NY, USA), pp. 214–227, ACM, 2018.

[6] C. Lattner, V. Adve, "LLVM: A Compilation Framework for Lifelong Program Analysis & Transformation," in *Proceedings of the International Symposium on Code Generation and Optimization: Feedback-directed and Runtime Optimization*, CGO '04, (Washington, DC, USA), pp. 75–, IEEE Computer Society, 2004.

[7] A. D. Malony, S. Biersdorff, W. Spear, S. Mayanglambam, "An Experimental Approach to Performance Measurement of Heterogeneous Parallel Applications Using CUDA," in *Proceedings of the 24th ACM International Conference on Supercomputing*, ICS '10, (New York, NY, USA), pp. 127–136, ACM, 2010.

[8] M. Stephenson, S. Hari, S. Kumar, Y. Lee, E. Ebrahimi, D. R. Johnson, D. Nellans, M. O'Connor, and S. W. Keckler, "Flexible Software Profiling of GPU Architectures," *SIGARCH Comput. Archit. News*, vol. 43, pp. 185–197, June 2015.

[9] R. Dietrich, R. Tschüter, "A generic infrastructure for OpenCL performance analysis," in *2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, vol. 1, pp. 334–341, Sept 2015.

[10] M. Haidl, B. Hagedorn, S. Gorlatch, "Programming GPUs with C++14 and Just-In-Time Compilation," in *Parallel Computing: On the Road to Exascale, Proceedings of the International Conference on Parallel Computing, ParCo 2015, 1-4 September 2015, Edinburgh, Scotland, UK*, pp. 247–256, 2015.

[11] F. Feldmann, B. Hagemann, L. Ganzer, M. Panfilov, "Numerical simulation of hydrodynamic and gas mixing processes in underground hydrogen storages," *Environmental Earth Sciences*, vol. 75, p. 1165, Aug 2016.

[12] Daniel J. Duffy, *Finite Difference methods in financial engineering: a Partial Differential Equation approach*. John Wiley & Sons, 2013.

[13] I. Rucker, W. Ressel, "A numerical drainage model to simulate infiltration into porous pavements for higher road safety," in *17. Internationales Stuttgarter Symposium*, (Wiesbaden), pp. 1293–1303, Springer Fachmedien Wiesbaden, 2017.

[14] Anthony W. Leung, *Systems of nonlinear partial differential equations: applications to biology and engineering*, vol. 49. Springer Science & Business Media, 2013.

[15] M. S. Alnæs, J. Blechta, J. Hake, *et al.*, "The FEniCS project version 1.5," *Archive of Numerical Software*, vol. 3, no. 100, pp. 9–23, 2015.

[16] A. Logg, K. B. Ølgaard, M. E. Rognes, and G. N. Wells, "FFC: the FEniCS form compiler," in *Automated Solution of Differential Equations by the Finite Element Method*, pp. 227–238, Springer, 2012.

[17] G. R. Markall, F. Rathgeber, L. Mitchell, *et al.*, "Performance-Portable Finite Element Assembly Using PyOP2 and FEniCS," in *Supercomputing*, (Berlin, Heidelberg), pp. 279–289, Springer Berlin Heidelberg, 2013.

[18] F. Rathgeber, D. A. Ham, L. Mitchell, *et al.*, "Firedrake: automating the finite element method by composing abstractions," *ACM Transactions on Mathematical Software (TOMS)*, vol. 43, no. 3, p. 24, 2017.

[19] D. L. Brown, W. D. Henshaw, and D. J. Quinlan, "Overture: An object-oriented framework for solving partial differential equations," in *International Conference on Computing in Object-Oriented Parallel Environments*, pp. 177–184, Springer, 1997.

[20] C. Prud'Homme, V. Chabannes, V. Doyeux, *et al.*, "Advances in Feel++: a domain specific embedded language in C++ for partial differential equations," in *Eccomas' 12-European Congress on Computational Methods in Applied Sciences and Engineering*, 2012.

[21] P. Bastian, M. Blatt, C. Engwer, *et al.*, "The distributed and unified numerics environment (DUNE)," in *Proc. of the 19th Symposium on Simulation Technique in Hannover*, vol. 123, 2006.

[22] P. Bastian, M. Blatt, A. Dedner, *et al.*, "A generic grid interface for parallel and adaptive scientific computing. Part I: abstract framework," *Computing*, vol. 82, no. 2-3, pp. 103–119, 2008.

[23] J. E. Stone, D. Gohara and G. Shi, "OpenCL: A Parallel Programming Standard for Heterogeneous Computing Systems," *IEEE Des. Test*, vol. 12, pp. 66–73, May 2010.

[24] J. Dongarra, K. London, S. Moore, P. Mucci and D. Terpstra, "Using PAPI for Hardware Performance Monitoring on Linux Systems," 08 2009.

[25] "CUPTI API." https://docs.nvidia.com/cuda/cupti/r_main.html. Accessed: 2018-07-14.

[26] P. Bastian, F. Heimann, and S. Marnach, "Generic implementation of finite element methods in the distributed and unified numerics environment (DUNE)," *Kybernetika*, vol. 46, no. 2, pp. 294–315, 2010.

[27] M. Blatt and P. Bastian, "The iterative solver template library," in *International Workshop on Applied Parallel Computing*, pp. 666–675, Springer, 2006.

[28] M. S. Alnæs, A. Logg, K. B. Ølgaard, M. E. Rognes, and G. Wells, "Unified Form Language: A Domain-specific Language for Weak Formulations of Partial Differential Equations," *ACM Trans. Math. Softw.*, vol. 40, pp. 9:1–9:37, Mar. 2014.

[29] M. Haidl, S. Moll, L. Klein, H. Sun, S. Hack, and S. Gorlatch, "PACXXv2+RV: An LLVM-based Portable High-Performance Programming Model," in *Proceedings of the Fourth Workshop on the LLVM Compiler Infrastructure in HPC*, LLVM-HPC, (New York, NY, USA), pp. 7:1–7:12, ACM, 2017.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 5, 123-131 (2019)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Green Inventory Supply Chain Model with Inflation under Permissible Delay in Finite Planning Horizon

Seema Saxena[*,1], Vikramjeet Singh[2], Rajesh Kumar Gupta[1], Nitin Kumar Mishra[1], Pushpinder Singh[3]

[1]*Department of Mathematics, Lovely Professional University, Phagwara 144411, Punjab, India*

[2]*Department of Mathematics, IKGPTU Campus,Batala, 143506, Punjab, India*

[3]*Department of Mathematics, Mata Gujri College, Fatehgarh Sahib, 140407, Punjab, India*

ARTICLE INFO

Article history:
Received: 12 July, 2019
Accepted: 19 August, 2019
Online: 24 September, 2019

Keywords:
Inflation
Supply chain
Credit term
Remanufacturing

ABSTRACT

Trade credit is an important cost reduction tool in the inventory management. The effect of trade credit is studied on the integrated system for sharing the cost benefits realized due to the permissible delay. Credit term factor is introduced to divide the cost benefits between the retailer and the supplier. The various costs in the inventory model are subjected to the same inflation rate. This research paper revisits EOQ model for remanufacturing process under green supply chain with the permissible delay available to the retailer. Numerical examples prove that the optimal re-ordering schedule exists and is unique. Also sensitivity analysis is performed on certain parameters to ascertain their logical implications.

1 Introduction

This paper is an extension of the work [1] originally presented in the 4th International Conference on Computing Science(ICCS),2018. This extended research work incorporates the green inventory concept. The harmful effects of the waste and outdated products are imposing a threat to the environment. In present scenario, world is facing pollution as a big hazard to mankind. Every organization is moving towards reducing and reusing the waste/imperfect goods. In this direction this EOQ model aims at a single stage remanufacturing process where the imperfect goods after screening are taken back by the supplier and are remanufactured. This is again transported to the retailer. Replenishment schedule is derived for the retailer and the supplier, considering green product life cycle and the time value of money. Various countries have adopted several measures for waste product management, reusing and remanufacturing programs. International standards as European Union's proposal for Waste Electrical and Electronic equipment (WEEE) introduces the concept of product design, to bring a decrease in the cost of disassembly and remanufacturing [2]. Product, if designed significantly reduces the cost of inspection, disassembly, repair, remanufacturing and recycling. In [3], the author first determine an optimal ordering size with recovery and remanufacturing. He studied the traditional EOQ model with continuous and deterministic demand and return. In [4], the authors probed the changes taking place in the market and in the manufacturing organizations. He studied the changes for the process of product design, with respect to the technology introduced for new material and new production methods, including tools and techniques altered for the manufacturing, inspection, reusing and remanufacturing. In [5], the author provided a comprehensive and immense knowledge for the remanufacturing process from the literature surveyed. The supplier and retailer are stationed far from each other so it is not possible for the supplier to send all perfect goods. Thus to be assure of the brand and quality, retailer screens the lot as it is received. In [6], the authors studied the remanufacturing of the imperfect items to maximize the total profit. The effect of deterioration is dominant and its consequences cannot be ignored while framing an EOQ model. Electronic goods, blood, fruits, grain products, alcohol are some of the deteriorating products. In [7], the authors were the first to study deterioration in an inventory model. Several researchers as [8], [9], [10]. [11] studied different patterns of deterioration in inventory models. In [12], the authors developed an inventory model for the stock dependent demand for deteriorating products under two level credit. In [13], the researcher studied a two-echelon supply chain model for deteriorating products under trade credit. They analyzed two models, one with demand being stock dependent and the second with demand being selling price dependent for the perishable products. In [14], the authors studied

[*]Corresponding Author Seema Saxena, Department of Mathematics, Lovely Professional University, Phagwara 144411, Punjab(India), 8591188250 & mseemamishra@gmail.com

an inventory model for the deteriorating products where the deterioration being time dependent. They studied permissible delay policy in their model. Partially backlogged shortages were als incorporated. In [15], the authors discusses the inventory model for perishable items where deterioration follows exponential distribution under permissible delay.
In conventional EOQ model, assumes that the retailer must pay for the goods as they arrive. But as practice, sometimes the supplier allows permissible delay to the retailer for the settlement of the goods. Numbers of research papers have been published with inventory models under trade credit. The rudest approach of an inventory model under trade credit was done by [16]. [17], developed an EOQ model under two level supply chain with trade credit for deteriorating products and demand being stock dependent. In [18],[19], the authors revisited the inventory models with time and stock dependent demand for imperfect and perishable products allowing trade credit. [20] studies the effect of sales effort on demand in a supply chain inventory model when permissible delay is allowed. He also included quantity discounts in his model with two level supply chain. Trade credit is widely used by the industries in U.S., China and Europe as discussed in [21],[22] and [23]. Suppliers give credit period to the retailer to postpone their payments, thereby increasing their participation and retaining the market. On the other hand, retailer can collect the revenue of the goods during the credit period and is encouraged to increase the order quantity.The concept of trade credit incorporated with the green inventory and weibull deterioration was studied by [24]. In [25], the authors analyses an inventory system with trade credit for imperfect goods and demand being quality dependent. He has considered two different approaches of trade credit in the supply chain model.

The formal EOQ Model establishing various results does not include the inflation for all the inventory related costs, whereas its effect should not be ignored. The first initiation for an inventory model under the time value of money was considered by [26]. In [27], the researchers analyzed an inventory model with the time value of money and weibull deterioration. In [28], the authors proposed a deteriorating EOQ model when demand is poisson and lead time is non zero. In [29], the authors developed a replenishment model for perishable products with return policy under inflation. In [30], the authors determine a reordering policy in an inventory model for deteriorating goods,partial backordering under inflation. In [1], the authors represented supply chain inventory model with inflation for time quadratic demand and deterioration. In [31], the authors developed an inventory model for deteriorating products considering inflation in a green supply chain for remanufacturing and recycling. Also a fuzzified model was developed by [32].

2 Assumptions and Terminology

2.1 Assumptions

The following assumptions are followed:

1. All inventory related cost are subjected to the same constant rate of inflation r.
2. The supplier offers the retailer credit period ρ to settle the account.
3. Finite replenishment and non zero lead time is considered.
4. The available inventory deteriorates with a constant fraction θ_1 in a finite planning horizon.
5. Single item single retailer and single supplier is considered in the supply chain.
6. A percentage p of products received in the lot are of imperfect quality.
7. The model is proposed in a finite planning horizon H, having I_{oi} units at the start of the cycle.

2.2 Terminology

1. The Opportunity cost $C_c(\$/unit/year)$ is same for the retailer and the supplier.
2. M_j^S is the permissible delay given by the supplier to the retailer for the settlement of the account in the synchronized model.
3. Setup cost for the supplier is S_s(\$/order).
4. The purchasing price for the supplier and the retailer is P_s(\$/unit) and P_o(\$/unit) with $P_s < P_o$.
5. In the synchronized model, the rate at which the additional cost is distributed by both the divisions is ρ.
6. Holding cost of the retailer and the suppler is H_o and H_s respectively.
7. Deterioration and Screening cost are DC and SC respectively.
8. The total inventory carried during a cycle is R_i^D in desynchronized model.
9. Total quantity in desynchronized model is Q^D.
10. In green inventory synchronized model, TC, REM and DSM are transportation cost, re-manufacturing cost and disassembly cost respectively.
11. In desynchronized model, the present estimate of the total cost for the retailer and the supplier is $PETC_r^D$ and $PETC_s^D$.

3 Model presentation and Analysis

The model is developed for two different scenarios:
Scenario 1 : Desynchronized Model

The retailer orders the quantity as per his convenience and the supplier follows the replenishment schedule of the retailer under inflation and green's technology of remanufacturing and recycling.
Scenario 2: Synchronized Model under permissible delay

The supplier offers credit period to the retailer to pay back the money for the lot received. Also shares the extra cost incurred by the retailer on account of following the replenishment schedule of the supplier with the order size increased. The present estimate of the total cost for the retailer and the supplier is calculated under inflation and green's technology of remanufacturing and recycling.

Model Description

The lot received at the beginning of the cycle has p percentage of imperfect quality goods. The complete lot undergoes screening process by the retailer and the defective goods are taken back by the supplier for remanufacturing and recycling. The cycle starts at t_i with I_{oi} units. The fall in the inventory is due to demand and deterioration till $t_i^{'}$.The level of inventory at this first stage of the process is $I_{1i}(t)$. At $t_i^{'}$, a decrease in demand occurs when the imperfect quality goods are taken back by the supplier. From $t_i^{'}$ to $t_i^{''}$, the fall in inventory is due to demand and deterioration. $I_{2i}(t)$ is the inventory for the second stage of the process. At $t_i^{''}$, an increase in the level of inventory is observed when the imperfect goods remanufactured are transported back to the retailer. Further at $t_i^{''}$, the inventory level $I_{3i}(t)$ declines due to demand and deterioration till it reaches zero. This behaviour is illustrated by figure 1.

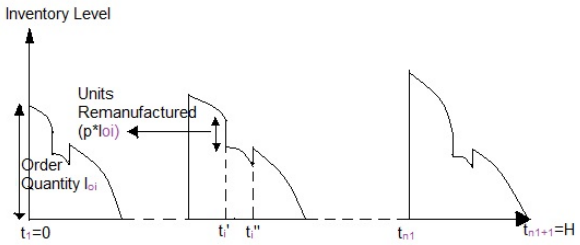


Figure 1: Graphical representation of Inventory Model

Scenario 1: Desynchronized Model

$$\begin{aligned}&\frac{d\left(I_{1i}(t)\right)}{dt}+\theta_1(t)I_{1i}(t)\\&=-f(t), t_i \le t\\&\le t_i^{'}, (i\\&=1,2,\ldots,n_1)\end{aligned} \tag{1}$$

and the boundary condition is $I_{1i}\left(t_i\right)=I_{oi}$ and $I_{1i}\left(t_i^{'}\right)=I_{si}$

$$\begin{aligned}&\frac{d\left(I_{2i}(t)\right)}{dt}+\theta_1(t)I_{2i}(t)\\&=-f(t), t_i^{'} \le t\\&\le t_i^{''}, (i\\&=1,2,\ldots,n_1)\end{aligned} \tag{2}$$

and the boundary condition is $I_{2i}\left(t_i^{'}\right)=I_{si}-p.I_{oi}$ and $I_{2i}\left(t_i^{''}\right)=I_{fi}$

$$\begin{aligned}&\frac{d\left(I_{3i}(t)\right)}{dt}+\theta_1(t)I_{3i}(t)\\&=-f(t), t_i^{''} \le t\\&\le t_{i+1}, (i\\&=1,2,\ldots,n_1)\end{aligned} \tag{3}$$

the boundary condition is $I_{3i}\left(t_i^{''}\right)=I_{fi}+p.I_{oi}$ and $I_{3i}\left(t_{i+1}\right)=0$
Solving these equations, we get

$$\begin{aligned}I_{\mathrm{oi}}=&\,a\left(t_{i+1}-t_i\right)+\left((\mathrm{a}\theta_1+b)/2\right).\left(t_{i+1}^2-t_i^2\right)+\left(\mathrm{b}\theta_1/3\right)\left(t_{i+1}^3-t_i^3\right)\\&-\left(\mathrm{p}\theta_1\left(t_{i+1}^{''}-t_i^{'}\right)+\theta_1 t_i\right).\left(a\left(t_{i+1}-t_i\right)+(b/2).\left(t_{i+1}^2-t_i^2\right)\right)\end{aligned} \tag{4}$$

$$\begin{aligned}I_{si}=&\,e^{-\theta_1.t_i^{'}}\left(1-p\left(-e^{-\theta_1\left(t_i-t_i^{'}\right)}+e^{-\theta_1\left(t_i-t_i^{''}\right)}\right)\right).\\&\int_{t_i}^{t_{i+1}}e^{\theta_1.u}.f(u)du-\int_{t_i}^{t_i^{'}}e^{\theta_1.\left(u-t_i^{'}\right)}.f(u)du\end{aligned} \tag{5}$$

$$\begin{aligned}I_{fi}=&\,I_{oi}.e^{\theta_1.\left(t_i-t_i^{''}\right)}-e^{-\theta_1 t_i^{''}}.\int_{t_i}^{t_i^{'}}e^{\theta_1.u}.f(u)du\\&-p.I_{oi}.e^{\theta_1.\left(t_i^{'}-t_i^{''}\right)}-e^{-\theta_1 t_i^{''}}.\int_{t_i^{'}}^{t_i^{''}}e^{\theta_1.u}.f(u)du\end{aligned} \tag{6}$$

substituting these values, the solution of the differential eq (1) is

$$\begin{aligned}I_{1i}(t)=&\,a\left(t_{i+1}-t\right)+\left((a.\theta_1+b)/2\right).\left(t_{i+1}^2-t^2\right)\\&+\left((b.\theta_1)/3\right).\left(t_{i+1}^3-t^3\right)+\left(p.\theta_1.t_i^{'}-p.\theta_1.t_i^{''}\right).\left(a\left(t_{i+1}\right.\right.\\&\left.\left.-t_i\right)+(b/2).\left(t_{i+1}^2-t_i^2\right)\right)-\theta_1.t.\left(a\left(t_{i+1}-t\right)\right.\\&\left.+(b/2).\left(t_{i+1}^2-t^2\right)\right), t_i \le t \le t_i^{'}, \quad \{i=1,2,\ldots,n_1\}.\end{aligned} \tag{7}$$

the solution of the differential eq (2) is $I_{2i}(t)=$
$(1-\theta_1.t).\left(a\left(t_{i+1}-t\right)+(b/2).\left(t_{i+1}^2-t^2\right)\right)+$
$\left((a.\theta_1)/2\right).\left(t_{i+1}^2-t^2\right)+\left((b.\theta_1)/3\right).\left(t_{i+1}^3-t^3\right)$
$-p.\left(\left((a.\theta_1)/2\right).\left(t_{i+1}^2-t_i^2\right)+\left((b.\theta_1)/3\right).\left(t_{i+1}^3-t_i^3\right)\right)-$

$$\begin{aligned}&p.\left(1+\theta_1.t_i^{''}-\theta_1.t-\theta_1*t_i\right).\left(a\left(t_{i+1}-t_i\right)+(b/2).\left(t_{i+1}^2-t_i^2\right)\right)\\&+p^2\left(\theta_1.t_i^{''}-\theta_1.t_i^{'}\right), t_i^{'} \le t \le t_i^{''}, \quad \{i=1,2,\ldots,n_1\}\end{aligned} \tag{8}$$

the solution of the differential eq (3) is

$$\begin{aligned}I_{3i}(t)=&\,(1-\theta_1.t).\left(a\left(t_{i+1}-t\right)+(b/2).\left(t_{i+1}^2-t^2\right)\right)+\\&\left(\left((a.\theta_1)/2\right).\left(t_{i+1}^2-t^2\right)+\left((b.\theta_1)/3\right).\left(t_{i+1}^3-t^3\right)\right), t_i^{''} \le t \le t_{i+1},\\&\{i=1,2,\ldots,n_1\}.\end{aligned} \tag{9}$$

Total inventory carried during the interval $[t_i, t_{i+1}]$ is

$$\begin{aligned}R_i^D=&\int_{t_i}^{t_i^{'}}I_{1i}(t)dt+\int_{t_i^{'}}^{t_i^{''}}I_{2i}(t)dt+\int_{t_i^{''}}^{t_{i+1}}I_{3i}(t)dt, \quad (i\\&=1,2,3,\ldots,n_1).\end{aligned} \tag{10}$$

Total quantity during the planning horizon =

$$Q^D=\sum_{i=1}^{n_1}I_{oi}^D \tag{11}$$

Retailer's total cost includes present estimate of the ordering cost, holding cost, deterioration cost, purchasing cost and screening cost. Present estimate of the retailer's total cost $PETC_r^D$ is

$$PETC_r^D=\sum_{i=1}^{i=n_1}\left(PEOC+PEHC+PEDC+PEPC+PESC\right)$$

$$\sum_{i=1}^{i=n_1} e^{(d-r)t_i}.n_1.O_r + H_o.$$
$$\int_{t_i}^{t_i'} e^{(d-r).t}.I_{1i}(t)dt + H_o.\int_{t_i'}^{t_i''} e^{(d-r).t}.I_{2i}(t)dt +$$
$$H_o.\int_{t_i''}^{t_{i+1}} e^{(d-r).t}.I_{3i}(t)dt + \text{DC}.\int_{t_i}^{t_i'} e^{(d-r).t}.\theta_1.I_{1i}(t)dt$$
$$+ \text{DC}.\int_{t_i'}^{t_i''} e^{(d-r).t}.\theta_1.I_{2i}(t)dt$$
$$+ \text{DC}.\int_{t_i''}^{t_{i+1}} e^{(d-r).t}.\theta_1.I_{3i}(t)dt + (P_o + SC).e^{(d-r).t_i}.I_{oi} \quad (12)$$

By determining the $t_i's$, the green's inventory model is optimized. Equation the first order derivative of the total cost to zero, the values are calculated.

$$\frac{\partial PETC_r^D(t_i;n_1)}{\partial t_i} = 0 \quad (13)$$

The optimal solution so obtained by solving the above equation is $n_1^{DO}, t_0, t_1^{DO}, t_2^{DO},\ldots,t_{n_1+1}^{DO} = H$,

With the existing replenishment schedule in the desynchronized model, the supplier determines his/her present estimate of the total cost by summing the setup cost,purchasing cost, holding cost, transportation cost, dissemmbly cost and the remaufacturing cost for the planning horizon H. The equation is as follows:

$$PETC_s^D =$$
$$n_1^{DO}.e^{(d-r)t_i}.S\,s + P_s.e^{(d-r)t_i}.I_{oi} + H_s.\int_{t_i'}^{t_i''} e^{(d-r)t}.p.I_{oi}\,dt +$$
$$(TC + DSM + REM).p.e^{(d-r)t}.I_{oi} \quad (14)$$

The optimized quantity to be ordered is:

$$Q^* = \sum_{i=1}^{n_1^{DO}} I_{oi}^{DO} \quad (15)$$

Scenario 2: Synchronized model with permissible delay

in the synchronized model supplier offers permissible delay to the retailer. On account of which the number of reordering cycle decreases,which further leads to a decrease in the setup cost of the supplier. With the new ordering plan, there is an increase in the cost for the retailer which the supplier compensates by distributing it through the parameter credit period rate.

For the new ordering plan n_2, the present estimate of the cost for the retailer is

$$PETC_r^S =$$
$$\sum_{j=1}^{j=n_2} (PEOC + PEHC + PEDC + PEPC + PESC) \quad (16)$$

The addition in the retailer's cost is given by:

$$PETC_r^S - PETC_r^D \quad (17)$$

The supplier compensates the retailer by adding the increase to his/her total cost and so the present estimate of the cost is

$$PETC_s^S =$$
$$n_2.e^{(d-r)t_j}.S\,s + P_s.e^{(d-r)t_j}.I_{oj} + H_s.\int_{t_j'}^{t_j''} e^{(d-r)t}.p.I_{oj}dt +$$
$$(TC + DSM + REM).p.e^{(d-r)t}.I_{oj} + PETC_r^S - PETC_r^D \quad (18)$$

The optimized value of the cost for the supplier $PETC_s^S$ in the synchronized model is determined in the same way as done for the retailer $PETC_r^D$ in the desynchronized model.

$n_2^{SO}, t_1^{SO}, t_2^{SO},\ldots,t_{n_2+1}^{SO} = H$ be the optimal solution for $PETC_s^S$.

The optimized ordering quantity in synchronized model is

$$Q^S = \sum_{j=1}^{n_2^{SO}} I_{oj}^{SO}$$

Equitable distribution of the extra cost incurred during synchronized model

For the retailer,the present estimate of the total cost in the synchronized model should be less than that in the desynchronized model, only then the new replenishment program is followed by the retailer.

$$PETC_r^D \geq PETC_r^S - \sum_{j=1}^{n_2^{SO}} e^{(d-r)*t_j}C_c.\rho\left(t_{j+1}^{SO} - t_j^{SO}\right)I_{oj}^{SO} \quad (19)$$

where ρ is the credit period rate which is the distributing factor for the extra cost at the retailer's end and the rate is same in all ordering intervals.

The credit period duration is given by:

$M_j^S = \rho\left(t_{j+1}^{SO} - t_j^{SO}\right)$

In the synchronized model,the supplier offers retailer credit for the minimum period.The credit period rate is kept minimum ρ_{min} to divide the benefits realized.The present estimate of his total cost in the synchronized model is then given by:

$$PETC_r^S - \sum_{j=1}^{n_2^{SO}} e^{(d-r)*t_j}C_c.\rho_{min}\left(t_{j+1}^{SO} - t_j^{SO}\right)I_{oj}^{SO} = PETC_r^D \quad (20)$$

which implies the value of ρ_{min} is:

$$\rho_{\min} = \frac{PETC_r^S - PETC_r^D}{\sum_{j=1}^{n_2^{SO}}.e^{(d-r)*t_j}.C_c\left(t_{j+1}^{SO} - t_j^{SO}\right)I_{oj}^{SO}} \quad (21)$$

As supplier bears the additional cost borne at the retailer's end, the credit period rate is maximum at his end ρ_{max}. His total cost in the synchronized model is given by the following equation:

$$n_2^{SO} S_s . e^{(d-r)*t_j} + \sum_{j=1}^{n_2^{SO}} P_s . e^{(d-r)*t_j} . I_{oj}^{SO} + \sum_{j=1}^{n_2^{SO}} . e^{(d-r)*t_j} . C_c .$$
$$\rho_{\max} \left(t_{j+1}^{SO} - t_j^{SO} \right) I_{oj}^{SO} = PETC_s^D \quad (22)$$

and ρ_{max} equals

$$\rho_{\max} = \frac{PETC_s^D - n_2^{SO} . e^{(d-r)*t_j} . S_s + \sum_{j=1}^{n_2^{SO}} P_s . e^{(d-r)*t_j} . I_{oj}^{SO}}{\sum_{j=1}^{n_2^{SO}} . e^{(d-r)*t_j} . C_c \left(t_{j+1}^{SO} - t_j^{SO} \right) I_{oj}^{SO}} \quad (23)$$

The additional expenditure due to the replenishment plan of the synchronized model is fractioned on the basis of $\bar{\rho}$, which is the average of the minimum and the maximum credit period rate. The present estimate of the net cost for both the retailer and the supplier is given by:

$$PETC_r^{SO\rho} = PETC_r^S \sum_{j=1}^{n_2^{SO}} . e^{(d-r)*t_j} . C_c \bar{\rho} \left(t_{j+1}^{SO} - t_j^{SO} \right) I_{oj}^{SO} \quad (24)$$

and $PETC_s^{SO\rho} =$

$$n_2^{SO} . e^{(d-r)*t_j} . S_s + \sum_{j=1}^{n_2^{SO}} P_s . e^{(d-r)*t_j} . I_{oj}^{SO} + \sum_{j=1}^{n_2^{SO}} . e^{(d-r)*t_j} . C_c . \bar{\rho} \left(t_{j+1}^{SO} - t_j^{SO} \right) I_{oj}^{SO} \quad (25)$$

4 Optimality Evaluation and Solution

The model is aimed at minimizing the total cost for the above cited two scenarios. The cycle time at which the cost is minimum is determined and the following theorem is verified with the help of the graphs.

Theorem 1: For any n_1,solution for the green supply chain EOQ model exists and is unique.

The convexity of the total cost can be seen through the following figures. Fig.2 plots the costs of retailer for the parameter SC = 4.5, and the value is minimum at the fourth cycle. The same is seen in Fig.3 for the supplier in the synchronized model where third cycle has a minimum value and the convexity of the cost is validated. The table values for the following figures is seen in Table 1 and Table 2 of the numerical example.

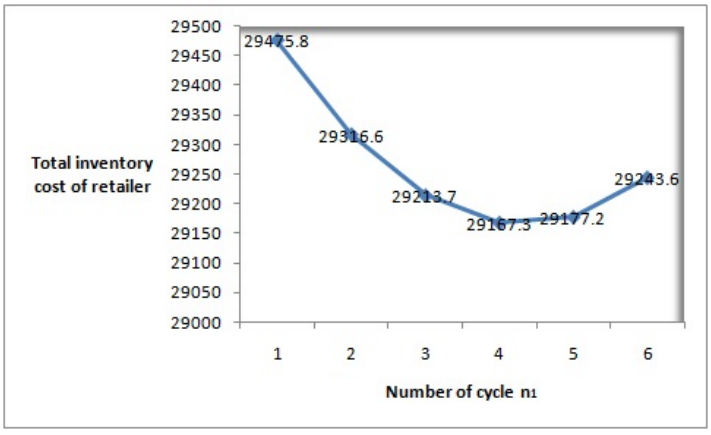


Figure 2: Convexity of total cost for retailer at SC=4.5 in desynchronized system

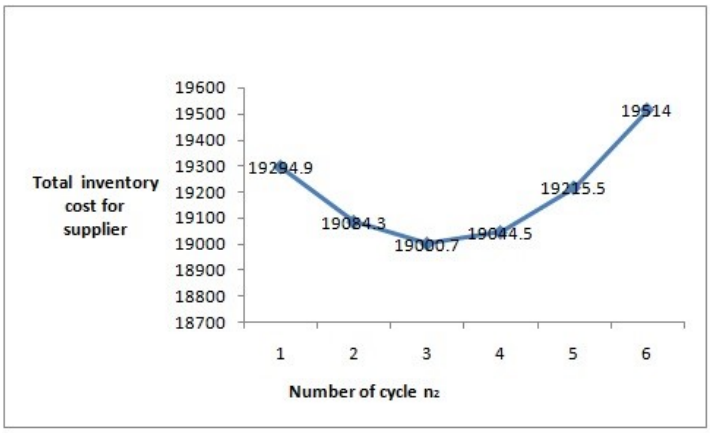


Figure 3: Convexity of total cost for supplier at SC=4.5 in synchronized system

5 Algorithm

1. All the parameters are assigned the hypothetical values.
2. In the desynchronized structure, the optimal reordering point for the retailer is determined.
   (a) Taking $t_1 = 0$, $n_1 = 1$ and $t_3 = H$. Initializing t_2, and calculating it from equation (13).
   (b) Taking $n_1 = 2$.
   (c) From the values of t_1 and t_2 the value of t_3 is calculated from equation (13).
   (d) The optimal values of $t_i's$ is determined for every n_1.
   (e) For $n_1 = 1$ and if $PETC_r^D(n_1) < PETC_r^D(n_1 + 1)$, then $PETC_r^D(n_1) = PETC_r^{DO}(n_1)$. Stop.
3. For $n_1 \geq 2$ and if $PETC_r^D(n_1) < PETC_r^D(n_1 - 1)$ and $PETC_r^D(n_1) < PETC_r^D(n_1 + 1)$, then $PETC_r^D(n_1) = PETC_r^{DO}(n_1)$ and stop or else let $n_1 = n_1 + 1$, and goto step 2(c).
4. $n_1^{DO} = n_1$ is the optimised cycle for both the retailer and the supplier in the desynchronized model.
5. $PETC_r^{DO}$, $PETC_s^{DO}$ and Q^{DO} are determined from the respective equations.
6. In synchronized model,the optimal cycle time n_2^{SO}, the present estimate of the cost for the retailer and the supplier $PETC_r^{SO}$ and $PETC_s^{SO}$ are calculated as done in steps from 2 to 4.
7. The values of $\rho_{\min}$, $\rho_{\max}$,$PETC_r^{SO\rho}$ and $PETC_s^{SO\rho}$ are determined from the respective equations.

5.1 Numerical Example

Hypothetical statistics is taken to validate the potency of the green inventory model.One principal parameter Screening Cost,SC is undertaken to observe the effects of changes in it on the optimal outcomes.

Example 1: a_1=20 items/year, b_1=5 items/year, $\theta_1 = 0.05$ items/year, $P_o = 4\$/item, \gamma = 0.01, t_1 = 0, O_r = 25$ \$/purchase order, $S_s = 30$ \$/arrangement, $H = 4$, $P_s = 0.01$ \$/item, $C_c = 2.5$ \$/item/year , $H_o = 100$ \$/item/year, d = 0.1, r = 0.05, DC = 70

$/item, CT = 0.1, TC=0.01 $/item, REM=0.01 $/item, DSM=0.01 $/item.

Table 1: Inflated cost of retailer in the desynchronized model under green's inventory system

SC	n_1		
	1	2	3
3	28568.6	28466.4	**28422.2**
3.75	28867.3	**28821**	28831.7
4.5	**29167.3**	29177.2	29243.6
SC	n_1		
	4	5	6
3	28568.6	28466.4	**28422.2**
3.75	28867.3	**28821**	28831.7
4.5	**29167.3**	29177.2	29243.6

The parameter of screening cost is further analyzed to study the impact of its changes on the green inventory model. Table 1 and Table 2 gives the total cost and the optimal ordering time period for both the retailer and the supplier in the desynchronized and the synchronized model respectively. Figure 2 and Figure 3 show the convexity of the cost.

Table 2: Inflated cost of supplier in the synchronized model under green's inventory system

SC	n_2		
	1	2	3
3	19707.5	19299.4	19020.9
3.75	19474.8	19165.2	18984
4.5	19294.9	19084.3	**19000.7**
SC	n_2		
	4	5	6
3	18872.2	**18853.3**	18964.4
3.75	**18931.1**	19006.8	19211
4.5	19044.5	19215.5	19514

Table 3 shows that the optimal number of replenishment schedules after the synchronization diminishes when the screening cost increases. This signifies that the supplier will be benefited is he/she reduces the number of ordering cycles. To add on, the percentage in the cost savings of the supplier also increases when the parameter SC increases, which is agreeable as the supplier compensates for the increase in the cost of the retailer in the synchronized system. Moreover same trend in the cost savings is seen for the retailer also. the percentage cost savings increases for the retailer when SC increases, which will assure the retailer's participation in the synchronized scheme.
Furthermore, the data also show that allowing for credit period, the optimal order quantity in the synchronized model is more than that as compared to the desynchronized model,thus raising the quantity ordered with reduced cycles.

Table 3: Percentage savings in cost for the retailer and the supplier in green's inventory system

Desynchronized system					
SC	$PETC_r^{DO}$	$PETC_s^{DO}$	n_1^{DO}	Q^{DO}	$\bar{\rho}$
3	28422.2	1914.67	6	118.912	0.728528
3.75	28821	1518.23	5	118.078	0.479235
4.5	29167.3	1195.08	4	117.79	0.273768

Synchronized system				
SC	$PETC_r^{SO\rho}$	$PETC_s^{SO\rho}$	n_2^{SO}	Q^{SO}
3	27730.2	1222.63	5	119.485
3.75	28131.3	828.518	4	119.485
4.5	28469.5	497.275	3	119.485

% Cost saving		
SC	$\frac{\Delta PETC}{PETC_r^{DO}}$	$\frac{\Delta PETC}{PETC_s^{DO}}$
3	2.43485	36.144
3.75	2.3931	45.4288
4.5	2.39243	58.3899

6 Theoretical aspects in green inventory model on trade credit

Hypothesis 1: Permissible delay is negatively associated with the inflation rate.

To deal with the changes in the financial costs that incur due to the changing rate of inflation, organisations frequently alter the policy of trade credit. There is a decrease in the credit period to cope with the inflation. This study has been made by several researchers.The following table shows that as the rate of inflation increases, there is a decrease in the credit period rate. The data also reveals that although inflation is anticipated yet there is a realization of cost savings for both the retailer and the supplier in the green inventory model.

Table 4: Effect of change in rate of inflation on credit period rate

r	$\bar{\rho}$	$\frac{\Delta PETC}{PETC_r^{DO}}$	$\frac{\Delta PETC}{PETC_s^{DO}}$
0.025	0.807762	2.30863	33.2966
0.05	0.728549	2.43486	36.1441
0.075	0.653612	2.55951	39.0887

Hypothesis 2: Permissible delay is negatively associated to the capital cost.

Supplier incurs extra cost due to permissible delay. However, both the retailer and the supplier at distinct rates can make financial investments. So with the increase in the capital cost, there is

a reduction in the credit period rate. Furthermore,the percentage savings in the cost for the retailer and the supplier is not associated with the alterations in the capital cost.

Table 5: Effect of change in rate of Capital cost on credit period rate

Cc	$\bar{\rho}$	$\frac{\Delta PETC}{PETC_r^{DO}}$	$\frac{\Delta PETC}{PETC_s^{DO}}$
1.5	1.45327	2.4252	36.1042
3	0.728548	2.43485	36.144
4.5	0.485699	2.43485	36.144

Hypothesis 3: Credit period rate is negatively associated with the deterioration.

When the inventory of finished goods is perishable, there is a decrease in the revenue collected due to deterioration. For the deteriorated goods the time duration for permissible delay also decreases.The hypothesis is validated by the numerical example and the following table. To add in the validation for the hypothesis [33],[34] and [35] too derived the same result through the mathematical model in their research work.

Table 6: Effect of change in rate of deterioration on credit period rate

θ_1	$\bar{\rho}$	$\frac{\Delta PETC}{PETC_r^{DO}}$	$\frac{\Delta PETC}{PETC_s^{DO}}$
0.0025	0.762206	2.26233	33.3653
0.005	0.728544	2.43483	36.144
0.0075	0.696866	2.70232	40.493

Hypothesis 4: Permissible delay is positively associated with the supplier's set up cost.

There is an increase in the credit period rate when the parameter Ss increases. This guarantees that the supplier prolongs the permissible delay period to attain the benefit from synchronization. A majority of works by [36], [37], [33] predicted through their work, the attribution of the firm's profitability with respect to the increase in the delayed time and the set up cost. There is percentage cost savings fruition for the supplier which enables him to extend credit period incentive to the retailer. The retailer too generates percentage savings in the cost,thereby motivational for him to accept the synchronization scheme.

Table 7: Effect of change in supplier's set up cost on credit period rate

Ss	$\bar{\rho}$	$\frac{\Delta PETC}{PETC_r^{DO}}$	$\frac{\Delta PETC}{PETC_s^{DO}}$
15	0.536511	1.38002	30.8952
30	0.728547	2.43485	36.144
45	0.931589	3.54378	39.4326

Hypothesis 5: The change in the optimal total cost of the retailer and the supplier is negatively associated with the parameter p, but permissible delay is positively associated with p.

With the high value of the inspection rate,the imperfect goods are quickly removed,thus reducing the cost of holding. The percentage savings is negatively associated as the goods removed decreases the time period of the revenue generation. The credit period increases as it takes time for the imperfect goods to be remanufactured and get again absorbed as demand in the greens inventory model, thus validating the hypothesis.

Table 8: Effect of change in the parameter p on credit period rate

p	$\bar{\rho}$	$\frac{\Delta PETC}{PETC_r^{DO}}$	$\frac{\Delta PETC}{PETC_s^{DO}}$
0.375	0.694987	2.78661	41.9272
0.75	0.737651	2.64678	39.2901
1.125	0.778312	2.49702	36.6757

6.1 Managerial Insights

The logical implications for the hypothesis derived are firm and clear through the numerical considered and also through the work done by various prominent researchers. First, the association of the delayed cash time period with the rate of inflation is shown. This suggest that, the longer delay time should be squeezed to short time period on account of increase in the rate of inflation. Although to reach to long term conclusions, further study on other parameters associated should be done. This hypothesis highlights the significance of the supply chain process in optimising the inventory and the inventory related cost. Next in the supply chain process, the parameters as the capital cost, deterioration, set-up cost and price are studied with the delayed time. Thus modelling a concept of payment time delay with the manangement of inventory can help the firms to grab the untapped gains.

6.2 Senstivity analysis

Sensitivity analysis is carried on the green inventory model analyzing whether the formulated model is influenced by the alterations in the input parameters.We analyze the consequence of the alterations in the factors against the changes in the total cost of the retailer, supplier and the credit period rate. Every parameter is altered by -50%, -25%, 25% and 50%, of the initial cost taken in Example 1.

The table 9 shows that the change in the cost of the retailer and the supplier increases as demand increases which is agreeable. Also there is a decrease in the percentage change in the credit period with the increase in the demand.

Table 9: Analysis on demand variable 'a'

value	$\frac{\Delta PETC_r^{SO\rho} * 100\%}{PETC_{rO}^{SO\rho}}$	$\frac{\Delta PETC_s^{SO\rho} * 100\%}{PETC_{sO}^{SO\rho}}$	$\frac{\Delta\bar{\rho} * 100\%}{\bar{\rho}_O}$
0.5	-30.0024	-29.83	2.8461
0.75	-15.36	-6.576	15.2355
1.25	15.3233	5.503	-8.733
1.5	30.6353	10.7837	-15.479

The analysis on the holding cost of the supplier indicates that with the increase in the cost value the change in the cost of the supplier increases and also the rate of credit period increases.This implies that the supplier's total cost is sensitive to the holding cost as shown in Table 10.

Table 10: Analysis on holding cost of supplier H_s

value	$\frac{\Delta PETC_r^{SO\rho} * 100\%}{PETC_{rO}^{SO\rho}}$	$\frac{\Delta PETC_s^{SO\rho} * 100\%}{PETC_{sO}^{SO\rho}}$	$\frac{\Delta\overline{\rho} * 100\%}{\overline{\rho}_O}$
35	0.5753	-13.003	-15.219
52.5	0.2850	-6.449	-7.548
87.5	-0.2848	6.452	7.551
105	-0.5697	12.905	15.105

With the analysis done on the ordering cost of the retailer as in Table 11, it is evident that when the ordering size increases the total cost of the retailer increases. But there is a decrease in the cost of the supplier as the number of cycles decreases which is the synchronized model.

Table 11: Analysis on the Ordering cost O_r

value	$\frac{\Delta PETC_r^{SO\rho} * 100\%}{PETC_{rO}^{SO\rho}}$	$\frac{\Delta PETC_s^{SO\rho} * 100\%}{PETC_{sO}^{SO\rho}}$	$\frac{\Delta\overline{\rho} * 100\%}{\overline{\rho}_O}$
12.5	-0.8777	17.69	19.444
18.75	-0.444	8.96	9.851
31.25	1.110	-24.81	-27.772
37.5	1.849	-47.07	-48.271

When the parameters P_s and REM are examined,the data is recorded in Table 12 and Table 13. It is seen that with the increase in the purchasing cost and the remanufacturing cost, percentage change in the the total cost of the supplier also increases which is apparent. The supplier's total cost is sensitive to these cost parameters. The credit period rate also increases wth the increase in these parameters. Although the retailer's cost is less responsive to these cost parameters. With the increase in the credit period the retailer can invests the amount to generate gains.

Table 12: Analysis on the Purchasing cost of the supplier P_s

value	$\frac{\Delta PETC_r^{SO\rho} * 100\%}{PETC_{rO}^{SO\rho}}$	$\frac{\Delta PETC_s^{SO\rho} * 100\%}{PETC_{sO}^{SO\rho}}$	$\frac{\Delta\overline{\rho} * 100\%}{\overline{\rho}_O}$
0.005	0.0093	-0.3012	-0.2619
0.0075	-0.000293	-0.0328	-0.00151
0.0125	0.000299	0.0327	0.00136
0.015	0.000594	0.0655	0.00283

Table 13: Analysis on the Remanufacturing cost of the supplier REM

value	$\frac{\Delta PETC_r^{SO\rho} * 100\%}{PETC_{rO}^{SO\rho}}$	$\frac{\Delta PETC_s^{SO\rho} * 100\%}{PETC_{sO}^{SO\rho}}$	$\frac{\Delta\overline{\rho} * 100\%}{\overline{\rho}_O}$
0.005	0.0102	-0.2518	-0.2946
0.0075	0.0002	-0.0138	-0.0162
0.0125	-0.0001	0.0137	0.0160
0.015	-0.0003	0.0276	0.0322

7 Conclusion

The research demonstrates a synchronized supply chain with permissible delay under inflation in a greens inventory system. The model shows that the organizations can adopt for the green technology favouring the environment and can still realize cost savings. The study shows that the cost under inflation in the synchronized model is not more than the cost in the desynchronized model. The primary characteristic of the synchronized model is the credit period rate. It divides the extra cost incurred to the retailer on account of accepting the new ordering plan by the supplier and the extra cost incurred to the supplier on account of green's technology of remanufacturing and recycling. The analysis of the model reveals that both the retailer and the supplier are able to generate financial benifits. Logical insights are proposed for the various parameters in this study.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment: The authors sincerely thank and appreciate the editor-in-chief and the anonymous reviewers who have considerably helped with their observations and insightful suggestions for the refinement of the paper.

References

[1] Vikramjeet Singh, Seema Saxena, Rajesh Kumar Gupta, Nitin Kumar Mishra, and Pushpinder Singh. A supply chain model with deteriorating items under inflation. In *2018 4th International Conference on Computing Sciences (ICCS)*, pages 119–125. IEEE, 2018.

[2] Michael W Toffel. End-of-life product recovery: Drivers, prior research, and future directions. In *Conference on European Electronics Take-back Legislation: Impacts on Business Strategy and Global Trade*. Citeseer, 2002.

[3] David A Schrady. A deterministic inventory model for reparable items. *Naval Research Logistics Quarterly*, 14(3):391–398, 1967.

[4] Maurice Bonney, Svetan Ratchev, and Idir Moualek. The changing relationship between production and inventory examined in a concurrent engineering context. *International Journal of Production Economics*, 81:243–254, 2003.

[5] Marisa MP de Brito. *Managing reverse logistics or reversing logistics management?* Number ERIM PhD Series; EPS-2004-035-LIS. 2004.

[6] Biswajit Sarkar, Waqas Ahmed, Seok-Beom Choi, and Muhammad Tayyab. Sustainable inventory management for environmental impact through partial backordering and multi-trade-credit-period. *Sustainability*, 10(12):4761, 2018.

[7] P. M. Ghare and G. F. Schrader. (1963) A Model for an Exponential Decaying Inventory. Journal of Industrial Engineering, 14, 238-243. *J. Ind. Eng.*, 14: 238–243, 1963.

[8] Richard P Covert and George C Philip. An eoq model for items with weibull distribution deterioration. *AIIE transactions*, 5(4):323–326, 1973.

[9] Fred Raafat. Survey of literature on continuously deteriorating inventory models. *Journal of the Operational Research society*, 42(1):27–37, 1991.

[10] Hui-Ming Wee. Economic production lot size model for deteriorating items with partial back-ordering. *Computers & Industrial Engineering*, 24(3):449–458, 1993.

[11] H-M Wee and C-J Chung c. Optimising replenishment policy for an integrated production inventory deteriorating model considering green component-value design and remanufacturing. *International Journal of Production Research*, 47 (5):1343–1368, 2009.

[12] Kun-Jen Chung. An algorithm for an inventory model with inventory-level-dependent demand rate. *Computers & Operations Research*, 30(9):1311–1317, 2003.

[13] M Rameswari and R Uthayakumar. An integrated inventory model for deteriorating items with price-dependent demand under two-level trade credit policy. *International Journal of Systems Science: Operations & Logistics*, 5 (3):253–267, 2018.

[14] Pushpinder Singh, Nitin Kumar Mishra, Vikramjeet Singh, and Seema Saxena. An eoq model of time quadratic and inventory dependent demand for deteriorated items with partially backlogged shortages under trade credit. In *AIP Conference Proceedings*, volume 1860, page 020037. AIP Publishing, 2017.

[15] Karuppuchamy Annadurai. Integrated inventory model for deteriorating items with price-dependent demand under quantity-dependent trade credit. *International Journal of Manufacturing Engineering*, 2013, 2013.

[16] Suresh Kumar Goyal. Economic order quantity under conditions of permissible delay in payments. *Journal of the operational research society*, 36(4):335–338, 1985.

[17] Kun-Jen Chung and Leopoldo Eduardo Cárdenas-Barrón. The simplified solution procedure for deteriorating items under stock-dependent demand and two-level trade credit in the supply chain management. *Applied Mathematical Modelling*, 37(7):4653–4660, 2013.

[18] Biswajit Sarkar. An eoq model with delay in payments and time varying deterioration rate. *Mathematical and Computer Modelling*, 55(3-4):367–377, 2012.

[19] T Sarkar, SK Ghosh, and KS Chaudhuri. An optimal inventory replenishment policy for a deteriorating item with time-quadratic demand and time-dependent partial backlogging with shortages in all cycles. *Applied Mathematics and Computation*, 218(18):9147–9155, 2012.

[20] Zhihong Wang and Shaofeng Liu. Supply chain coordination under trade credit and quantity discount with sales effort effects. *Mathematical Problems in Engineering*, 2018, 2018.

[21] Tsung-Te Lin and Jian-Hsin Chou. Trade credit and bank loan: Evidence from chinese firms. *International Review of Economics & Finance*, 36:17–29, 2015.

[22] Yong-Wu Zhou, Zong-Liang Wen, and Xiaoli Wu. A single-period inventory and payment model with partial trade credit. *Computers & Industrial Engineering*, 90:132–145, 2015.

[23] Brahima Coulibaly, Horacio Sapriza, and Andrei Zlate. Financial frictions, trade credit, and the 2008–09 global financial crisis. *International Review of Economics & Finance*, 26:25–38, 2013.

[24] Vikramjeet Singh, Nitin Kumar Mishra, Sanjay Mishra, Pushpinder Singh, and Seema Saxena. A green supply chain model for time quadratic inventory dependent demand and partially backlogging with weibull deterioration under the finite horizon. In *AIP Conference Proceedings*, volume 2080, page 060002. AIP Publishing, 2019.

[25] Brojeswar Pal. Optimal production model with quality sensitive market demand, partial backlogging and permissible delay in payment. *RAIRO-Operations Research*, 52(2):499–512, 2018.

[26] JA Buzacott. Economic order quantities with inflation. *Operational research quarterly*, pages 553–558, 1975.

[27] Sh-Tyan Lo, Hui-Ming Wee, and Wen-Chang Huang. An integrated production-inventory model with imperfect production processes and weibull distribution deterioration under inflation. *International Journal of Production Economics*, 106(1):248–260, 2007.

[28] Mohammadmahdi Alizadeh, Hamidreza Eskandari, S Mehdi Sajadifar, and Christopher D Geiger. Analyzing a stochastic inventory system for deteriorating items with stochastic lead time using simulation modeling. In *Proceedings of the Winter Simulation Conference*, pages 1650–1662. Winter Simulation Conference, 2011.

[29] Maryam Ghoreishi, Gerhard-Wilhelm Weber, and Abolfazl Mirzazadeh. An inventory model for non-instantaneous deteriorating items with partial backlogging, permissible delay in payments, inflation-and selling price-dependent demand and customer returns. *Annals of Operations Research*, 226(1):221–238, 2015.

[30] Vikramjeet Singh, Seema Saxena, Pushpinder Singh, and Nitin Kumar Mishra. Replenishment policy for an inventory model under inflation. In *AIP Conference Proceedings*, volume 1860, page 020035. AIP Publishing, 2017.

[31] Smita Rani, Rashid Ali, and Anchal Agarwal. Green supply chain inventory model for deteriorating items with variable demand under inflation. *International Journal of Business Forecasting and Marketing Intelligence*, 3(1):50–77, 2017.

[32] Pushpinder Singh, Nitin Kumar Mishra, Manoj Kumar, Seema Saxena, and Vikramjeet Singh. Risk analysis of flood disaster based on similarity measures in picture fuzzy environment. *Afrika Matematika*, 29(7-8):1019–1038, 2018.

[33] Daniel Seifert, Ralf W Seifert, and Olov HD Isaksson. A test of inventory models with permissible delay in payment. *International Journal of Production Research*, 55(4):1117–1128, 2017.

[34] Abubakar Musa and Babangida Sani. Inventory ordering policies of delayed deteriorating items under permissible delay in payments. *International Journal of Production Economics*, 136(1):75–83, 2012.

[35] Xiping Song and Xiaoqiang Cai. On optimal payment time for a retailer under permitted delay of payment by the wholesaler. *International Journal of Production Economics*, 103(1):246–251, 2006.

[36] YW Zhou and D Zhou. Determination of the optimal trade credit policy: a supplier-stackelberg model. *Journal of the Operational Research Society*, 64 (7):1030–1048, 2013.

[37] Jianwen Luo and Qinhong Zhang. Trade credit: A new mechanism to coordinate supply chain. *Operations Research Letters*, 40(5):378–384, 2012.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 126-133 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

The Potential of Ocean Current as Electrical Power Sources Alternatives in Karimunjawa Islands Indonesia

Aris Ismanto[1,2 *], Dwi Haryo Ismunarti[1], Denny Nugroho Sugianto [1,2], Siti Maisyarah[1], Petrus Subardjo[1,] Agus Anugroho Dwi Suryoputro[1] , Hendry Siagian[1]

[1] *Oceanography Department, Faculty of Fisheries and Marine Science, Diponegoro University, Semarang, Indonesia 50275*

[2]*PUI-CoREM (Center of Excellence for Science and Technology - Center for Coastal Disaster Mitigation and Rehabilitation Studies), Diponegoro University, Indonesia*

ARTICLE INFO

Article history:
Received: 22 July, 2019
Accepted: 10 November, 2019
Online: 25 November, 2019

Keywords:
Renewable
Energy
Current
Model
Karimunjawa

ABSTRACT

Electrical energy shortage and expensive basic electricity costs are one of the problems that are occurred in Karimunjawa Island, Indonesia. The purpose of this research is to identify the potential of ocean currents energy as one of the alternatives to reduce the problem related to energy needs and to predict the electrical energy that can be obtained. The physical phenomena movement of ocean currents is made by using the 2-Dimensional hydrodynamic mathematical model. ADCP Multicell Argonaut-XR type is used to measure the speed of ocean currents to validate the model result. Tidal, wind, coordinate and bathymetry data are all the variables that are in the mathematical model of ocean current movement. The potential of electrical energy is determined by the value of power density which is calculated based on the speed ocean current model. The result of this research is the map of the potential distribution of ocean currents in Karimunjawa waters.

1. Introduction

Indonesia was a maritime country which is the largest archipelago with an area of 5,193,252 km^2, two-thirds of Indonesia's territory is an ocean, which is around 3,288,683 km^2 [1] and the whole island is about 17,508 islands [2]. Karimunjawa Islands situated on the Java Sea along of coordinate 5 '40 "- 5' 57" SL and 110 '4 "- 10' 40" EL. The area is around 111,625 ha, consisting of 7,033 ha of land and 104,592 ha of water. There are 5 inhabited islands out of 27 islands in Karimunjawa namley Karimunjawa, Kemujan, Parang, Nyamuk and Genting island [3]. One of the problems in the Karimunjawa islands is electricity. The occurrence of an electricity crisis in the form of service time duration is only lights up 6 hours each day and basic electricity cost are expensive [4]. Since the enactment of Government Regulation Number of 12, 2012 concerning retail cost and consumers of specific fuel users, diesel power plants sources (PLTD) in Karimunjawa must used industrial diesel at cost of IDR 14,700 each liter. The regional government spend money IDR 12 billion each year to welshed supply electricity [5].

*Ismanto, et al., Oceanography Department, Diponegoro University, +6285292419995. aris.ismanto@gmail.com

According to [6,7], the sea potential being a course of energy is very high. The stored energies the potential of wave energy, tidal range, ocean currents, ocean thermal & salinity gradient energy, and subsea geothermal energy. According to the Indonesian Marine Energy Association (ASELI), the total of marine resources in Indonesia reaches 727,000 MW [8].

Research related to ocean currents as a potential source of electrical energy has been conducted since 2006 in East Nusa Tenggara, i.e. in Lombok Strait, Alas Strait, Nusa Penida Strait, Flores Strait, Pantar Strait [9] and Larantuka waters, East Flores [10]. The results of research in Larantuka waters show that ocean currents have great potential to be developed as an energy source of the power plant. The characteristics of ocean currents are generated by tidal currents. The morphological condition of the waters in the presence of Gonzales strait causes an optimum current speed. Forecast that 96% of ocean currents in Larantuka waters has the potential to be developed as a power plant. The average power density generated is 6271,75 watts [11].

Some countries have managed to harness ocean current and tidal energy as the source of power plants, from a generator's turbine prototype to the commercial-scale turbines with a capacity

of 1.2 MW / turbine. Marine current energy plants have been built in Scotland, Sweden, France, Norway, England, Northern Ireland, Australia, Italy, South Korea, and the United States [12].

Ocean currents are horizontal motions of water masses so they have kinetic energy that can drive the rotor or turbine power plant as well as on wind power plants. An effort to explore a source of energy unconventional from the ocean currents begins with the mapping of the ocean currents and the potential energy. The movement of ocean currents that swept across the ocean currents in the water line data based observations and takes along time and is quite expensive in measuring data. An alternative approach mathematical models to study how the pattern of the current movement in the water line. The model is a prototype or imitation of real situation [13]. Over the past three decades the development of mathematical models has been rapidly driven by the development of numerical analysis, computer technology, and visualization techniques. The acceptance and confidence level of the mathematical model on hydrodynamics is currently very high, both in research, engineering and industry [14].

The Mathematical equation formulas that describe the physical phenomena of current movements include 2-D mathematical model in hydrodynamics. As for the force that can affect the movement of the current model is the tidal forces on the surface. Variables that will be used in the mathematical model of the movement of ocean currents conducted are tidal, wind coordinate and bathymetry data. Energy potential can be determined by the power density calculated based on the value of the current speed obtained from the ocean current model. Energgy potential was presented in the map of potential energy distribution.

The purpose of this research are carried out in developing studies of renewable energy sources of ocean current in the Karimunjawa Islands. Ocean current energy is renewable energy that can be used as potential energy in the coastal areas, especially in small islands [15]. Ocean current pattern mapping is the first step to explore non-conventional energy sources of ocean currents. Information on current movement patterns can also be used for various purposes such as consideration in the construction of harbor docks, onshore and offshore structures (drilling rig and pipes to be installed on the seafloor), aquaculture and site selection most likely for development of current power generation [13].

2. Materials and Methods

2.1. The research area

Quantitative method as the main method used in this study, the resulting data includes quantitative data as the main and supporting data. quantitative data in the form of data figures that will be analyzed according to the method [16]. there are 2 types of data used in this study, namely primary and secondary data. Data including secondary data is wind data for 10 years (2004-2014) with hourly records obtained from BMKG Semarang. Data obtained from DISHIDROS are coordinate data and bathymetry data. Secondary data is used to build 2-D mathematical model of current speed and direction of the current. Water flow data throughout the year are generated from 2-D mathematical modeling.

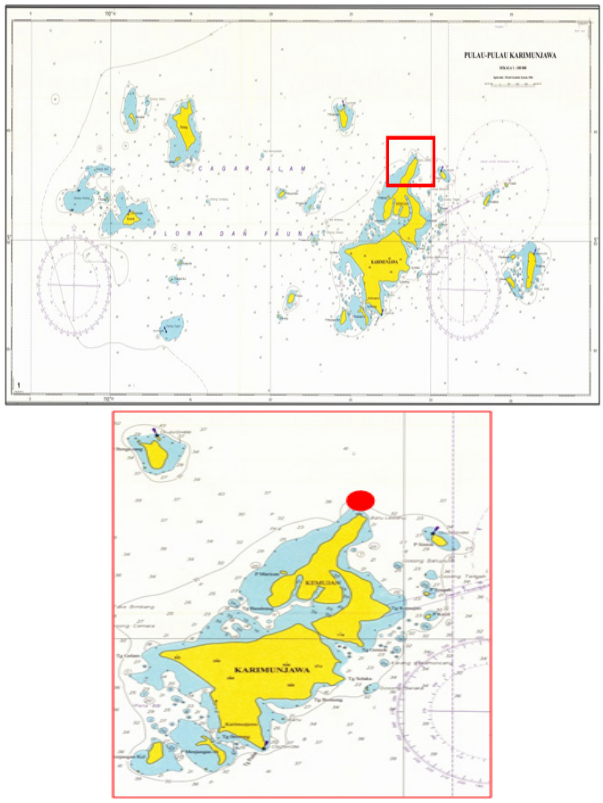


Figure 1. Study area and observation station [17].

Primary data measurement are tidal elevation and the current speed & direction data. Field data obtained are used to test the accuracy of the model. The current speed was measured using an ADCP Multi-Cell Argonaut-XR type device with wave length sensor beam 0.75 MHz and autonomous multi-cell system. Water depth is about 13 meters, while the total depth is 12 meters. Current speed is observed at eight points depth (d) or layers, i.e. at a depths of 1.5 meters, 3 meters, 4.5 meters, 6 meters, 7.5 meters, 9 meters, 10.5 meters, and 12 meters. Current speed measurements were carried out for 5 x 24 hours with measuring intervals at each layer and an average layer of 600 seconds. The obtained maximum current speed in the depths between 1.5 meters – 3 meters is 1.02 m/s. The measurement station point was determined based on the maximum current speed of the simulated numerical model at the same time as the field measurement [17].

2.2. Mathematical Model on Hydrodynamics

Ocean currents are the horizontal movement of seawater mass caused by the driving force acting on seawater. The main movement of water were caused by tides or the movements of sea levels due to the gravity [18-19], The current is also generated by the wind that blows on the surface of the water that transfers momentum and energy to the surface layer of water. Wind stress creates a tangential force which then moves the water in the direction of the wind.

The processes that occur in the fluid at sea are complex. The process described in a mathematical equations is called a mathematical model. Mathematical modeling on hydrodynamics to gain an understanding of ocean hydrodynamics. The goal is to study the mechanical variables of the fluid systems at sea. Utilization of models among others for tidal prediction, water level

elevation, current pattern, nutrient dispersion or pollutants, sedimentation and erosion.

The current movement patterns have vertical dimensions (depth) and horizontal dimensions. In general, the depth dimension is much smaller than the horizontal dimension so that the model that we used is a two-dimensional horizontal model (2DH). The mathematical equations that references the 2DH model is a continuity and momentum equation with simplicity: The mean value of depth is considered to represented parameters values that fluctuate along with the depth of flow. The speed and accelerationin the vertical direction are considered to be of small value, so they are ignored. the hydrostatic pressure distribution that applied together with the depth and the slope of base int both small horizontal directions.

The average value of the hydrodynamic variable at all depth is obtained by integrating it into the flow depth: $u = \frac{1}{h}\int_{z0}^{zt} u\, dz$. According to the study of Kowalik and Murty [20] the continuity and momentum equations for two DH current are:

The continuity equation :

$$\frac{\partial h}{\partial t} + \frac{\partial Uh}{\partial x} + \frac{\partial Vh}{\partial y} = 0 \quad (1)$$

The momentum equation in the X-axis direction:

$$\frac{\partial Uh}{\partial t} + \frac{\partial\left(U^2h + {}^1/_2 gh^2\right)}{\partial x} + \frac{\partial UVh}{\partial y} = gh\left(S_{ox} - S_{fx}\right) + \frac{\partial}{\partial x}\left(\frac{h}{\rho}T_{xx}\right) + \frac{\partial}{\partial y}\left(\frac{h}{\rho}T_{yx}\right) \quad (2)$$

The momentum equation in the y-axis direction :

$$\frac{\partial Uh}{\partial t} + \frac{\partial UVh}{\partial x} + \frac{\partial\left(U^2h + {}^1/_2 gh^2\right)}{\partial y} = gh\left(S_{oy} - S_{fy}\right) + \frac{\partial}{\partial x}\left(\frac{h}{\rho}T_{xy}\right) + \frac{\partial}{\partial y}\left(\frac{h}{\rho}T_{yy}\right) \quad (3)$$

Where:

x = distance in x-direction (direction of longitudinal flow), m

u = x-directional horizontal speed, m/s

y = distance in the y-direction (lateral flow direction), m

v = horizontal speed of the y-direction, m/s

t = time, s

h = water depth, m,

r = density, kg/m3

S_{ox} & S_{oy} is the basic slope of the x and y directions.

S_{fx}& S_{fy} is the slope of energy lines in the x and y-direction.

T_{ij} is the mean shear stress to depth

g = acceleration of gravity, $m/s2$

The slope of the energy line is calculated by the Manning or Chezy equation:

The Manning equation, i.e. $S_{fx} = \frac{n^2 U\sqrt{U^2+V^2}}{h^{\frac{4}{3}}}$ and $S_y = \frac{n^2 V\sqrt{U^2+V^2}}{h^{\frac{4}{3}}}$ (4)

The Chezy equation, i.e. : $S_{fx} = \frac{U\sqrt{U^2+V^2}}{C^2 h}$ and $S_y = \frac{V\sqrt{U^2+V^2}}{C^2 h}$ (5)

The basic friction coefficient is expressed by $\tau_x^b = \rho_w g \frac{U\sqrt{U^2+V^2}}{C^2}$ (6)

The basic coarse coefficient is $C = \frac{1}{n}h^{\frac{1}{6}}$ (7)

2.3. Ocean Current Energy as A Power Plant

Utilization of ocean currents as a source of electrical energy has same principle as wind power. Kinetic energy from ocean currents is generated from water that moves through water turbines. Compared to wind turbines, currents turbine will produce greater energy at the same speed. This occurs because the density of seawater is much higher than the density of air [21].

Current-generated power plants almost do not produce pollution. Current power generation also has no impact on current flows or sediment transfers as it affects only the pattern of ocean currents at the site. Slow turbine blade speeds are harmless to marine organisms and do not produce underwater sound pollution compared to ship propellers so it will cause into the small ecology [22]. In addition, the current has a tendency phase along with the rise and fall of the tides so that the power generated is easy to predict [23].

The energy potential is assumed in terms of power density which is proportional to its speed. The power assemblies generated by ocean currents can be calculated using the following equation: $P = \frac{1}{2}\rho A V^3$, where ρ is the density of seawater (kg/m^3); A is the area of turbine cross-section used (m^2), and V is the speed of ocean currents (m/s) [24]. Seawater density is assumed to be homogeneous with a value of ρ = 1.025 kg/m^3 [21].

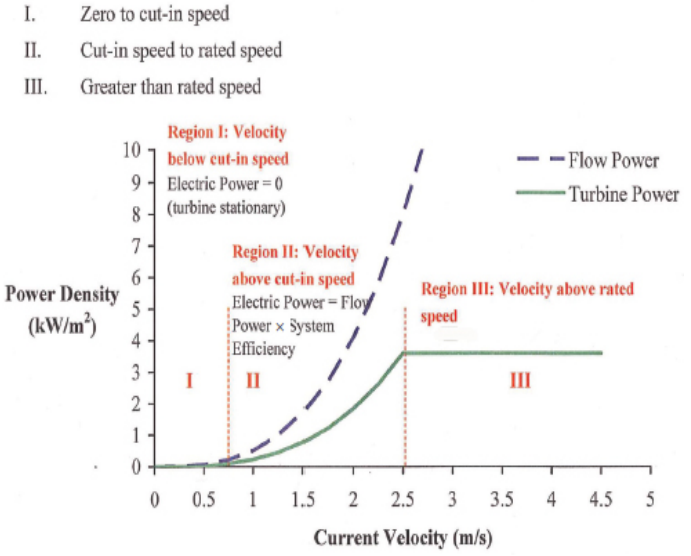


Figure 2. Relationship of power density and current speed; (I) Current speed less than cut-in speed; (II) Current speed between cut-in speed and rated speed; (III) Current speed above rated speed [21].

The statistical approach used is the potential probability analysis [25] Turbine boundary conditions used in the normal distribution are the Helix Gorlov turbine that is 0.5 m/s, while the normal distribution equation is [26] :

$$Z = \frac{\bar{x}-0.5}{S_x} = \frac{\bar{x}-0.5}{S/\sqrt{n}} \quad (8)$$

Turbine efficiency varies with current speed change. Cut-in speed is the minimum speed of current to be able to drive the turbine [27]. When the current speed is less than the cut-in speed, the kinetic energy to drive the turbine is insufficient and consequently, no power is produced. When the current speed is between the cut-in speed and rated speed, the power generated by the turbine can vary with the changes in current speed, which increases when the ocean current speed increases and decreases as the current speed decreases. Rated speed is the maximum speed that function as a barrier. Above that speed of the power generated by the turbine will be constant regardless of the current speed [23].

3. Results and Discussion

3.1. The validation of the result

The validation in this model used RMS (Root Mean Square) method [26], where RMS is a non-dimensional value to indicate a match two kinds of data. The set values of the model defined by b_1, b_2,..., b_n while the measurement define by a_1, a_2,..., a_nso we have:

$$X_{rms} = \left(\frac{\sqrt{\Delta x_1^2+\Delta x_2^2+\cdots+\Delta x_n^2}}{n}\right) \text{ where } \Delta xn=bn-an \quad (9)$$

The results validate the model results with field results indicate that the water level modeling results are 0.06 m (Figure 3) where the tolerance is 0,1 m, which means that the model has very good criteria. Karimunjawa waters generally have a tidal type of single daily tidal waters with a Formzahl value of 4.16 (F> 3) [28]. sinusoidal which is equal to water level. High current speed occurs when tide elevations lead to low tide. The direction of the current shows the direction of the current which has 2 directions (bi-directional current). Changes in current direction occur following changes in water level, during high tide the current direction is receding dominated by the direction 270º to 315º (west-northwest). At low tide towards the tide direction that dominates the direction is 40º to 70º (northeast) [17].

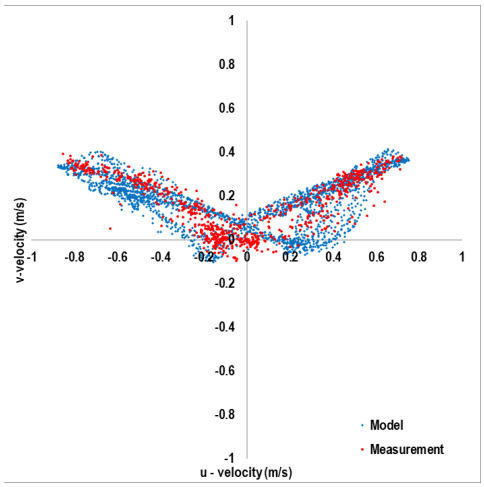
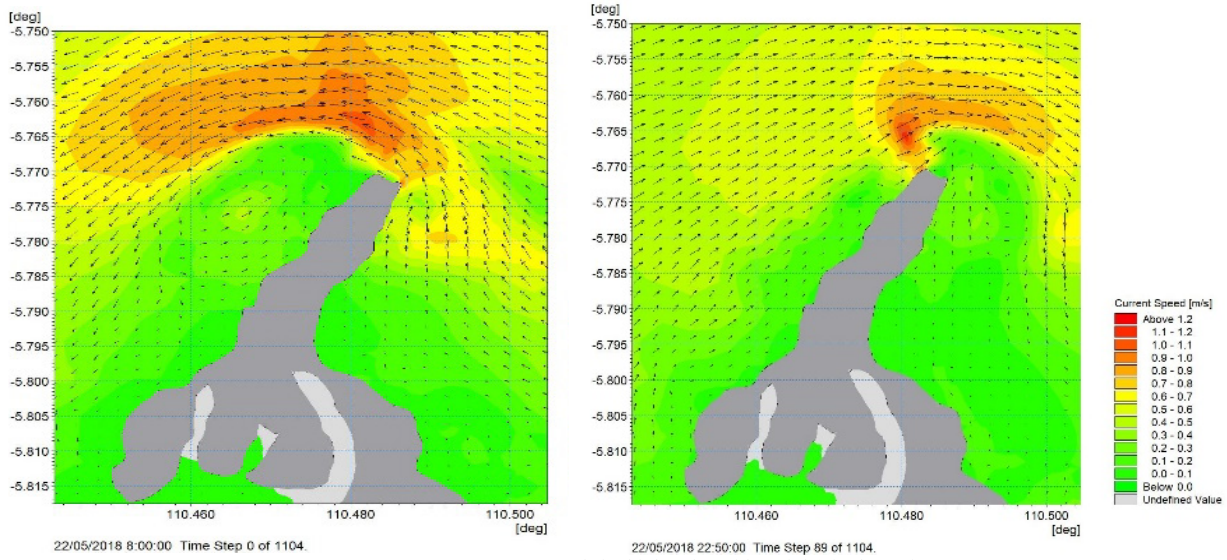


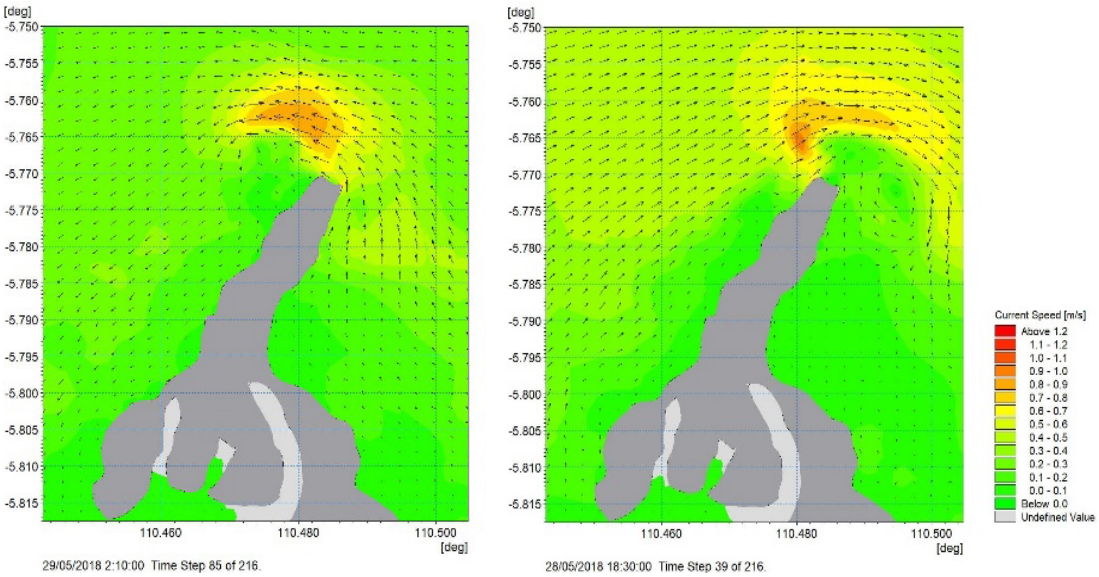
Figure 3. Validation result graphic between model and measurement vector velocity (Source: Data Processing, 2017)

Simulation of the ocean current numerical model in the depths between 1.5 meters – 3 meters was carried out in spring and neap tidal conditions. In this model tides is assumed has no influence of wind or basic friction. Full tides are formed due to the accumulation of lunar and sun attraction in a straight line. This statement is supported [29] who state that the earth's rotation and the relative position of the earth with respect to the moon and the sun affect the tides of the sea. The numerical simulation results are simplified by several complex parameters in the waters that are made into domains to facilitate the implementation of numerical simulations. Argues that to carry out simulations a system simplification is needed by maintaining the stability of the main components [30].

The model simulation is performed on high tides and low tides conditions (Figure 4 and 5). At the hight to low tides shows that the ocean current moves from NorthEast to SouthWest, while in the low to high tides the ocean currents moves from southwest to northeast where the highest current speed is 1.2 m/s. This is consistent with the statement from [31] tidal currents have the nature of moving in opposite directions or bi-directional. The direction of current during tide is usually the opposite of the direction of current when receding.



left), high to low tides (right) in the spring season



left) and high to low tides (right) in the neap season.

According to [32], tidal currents will affect daily flow patterns. Tidal currents generally flow from the coast towards the Indian Ocean when the tide and vice versa will flow from the Ocean to the coast at low tide. This current dominates the waters of Karimunjawa.

3.2. Current speed Result

The current conditions from the numerical model simulation is shown in all season condition (west season, east season, the transition 1 and the transition 2 seasons). Presentation of current patterns spatially is the method of presentation in which data speed and current direction are displayed based on location or specific coordinates. The peak month of each season is represented by the conditions that are taken in. From this figure (Figure 6,7,8 and 9), we can see the highest speed is 1.1 m/s in the West and East Season, while the highest speed in the transition 1 and 2 Season is 0.9 m/s.

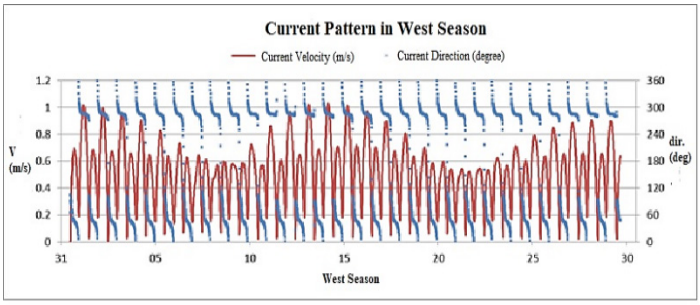


Figure 6 Current speed and direction graphic in west season

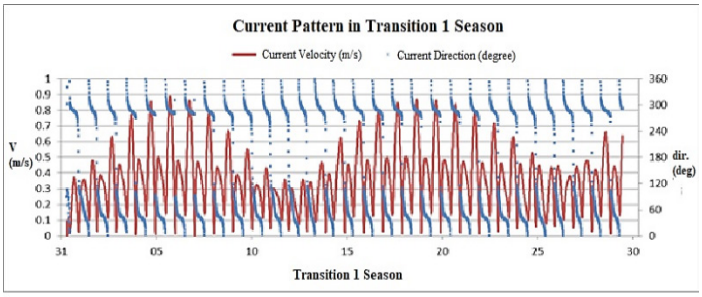


Figure 7. Current speed and direction graphic in the transition 1 season

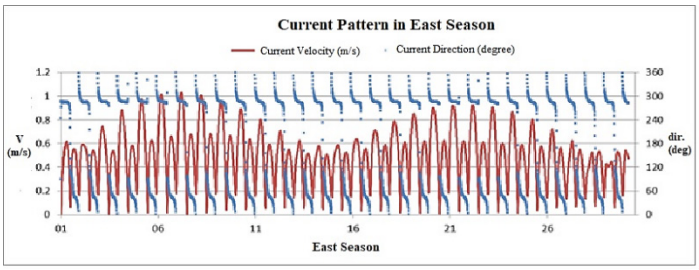


Figure 8. Current speed and direction graphic in east season

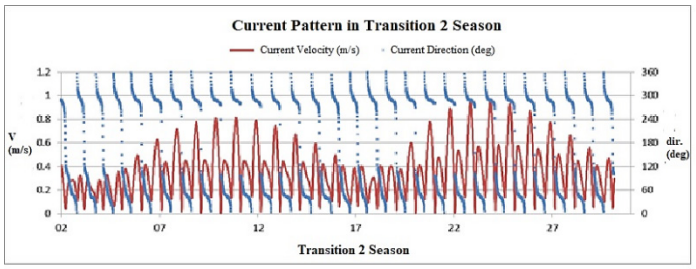


Figure 9. Current speed and direction graphic in transition 2 season

3.3. Analysis of Potential Energy

Overall current speed has a speed above the value of the turbine requirement limit (cut in speed) which is equal to 0.5 m / sec, in general, it can be seen that the current speed has the potential for ocean current energy in all seasons. The current speed from numerical model result is used to convert the estimation of electrical power calculations (watt) which is shown in all season conditions.

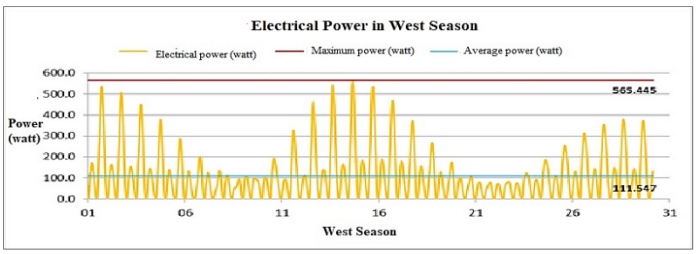


Figure 10. Potential power result (watts) in the west season

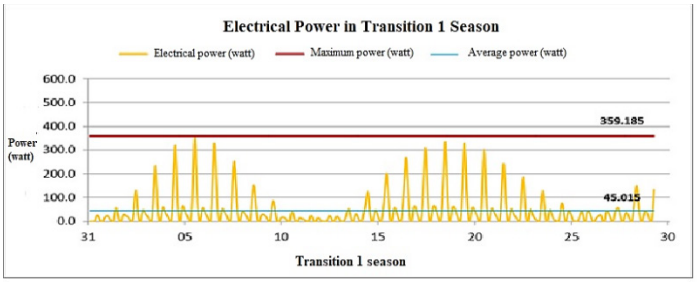


Figure 11. Potential power result (watts) in the transitional 1 season

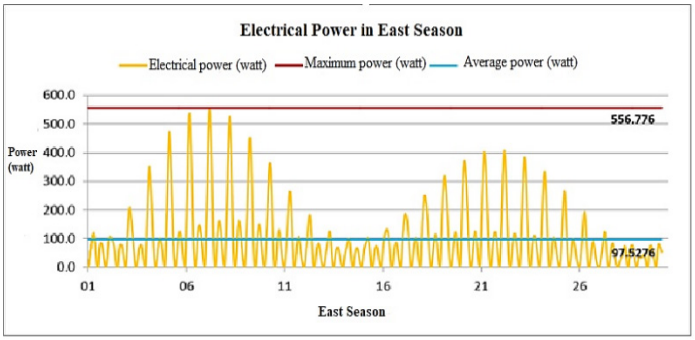


Figure 12. Potential power result (watts) in the east season

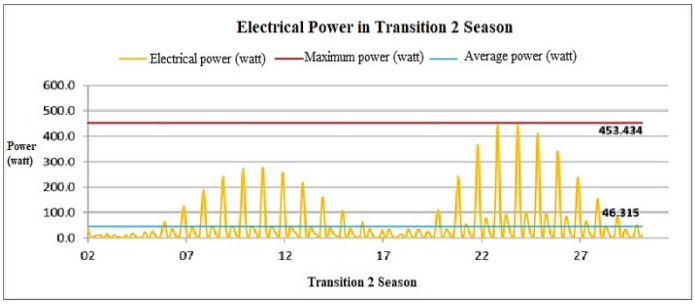


Figure 13. Potential power result (watts) in transitional 2 season

After we know the optimal potential of each season we can decide where will be the potential location of the ocean current power station. The highest potential Electrical Power in West season is 565.445 Watt, at the Transition 1 is 359.85 Watt, in the East Season is 556.76 Watt while in the transition 2 is 453.34 Watt. The analysis results are represented in Figures 10, 11, 12, and 13.

In the East Season the average of electrical power is 97.526 Watt and in the transition season is 46.315 Watt, while in the west season the highest average of electrical power obtained is 111.5 watts and the lowest average of electrical power is in the transition season 1 is 45.01 watts. We can see that the Electrical Power is really depended on the value of the ocean current speed (see Figure 10, 11, 12 and figure 13). The higher the speed value, the higher the electrical energy produced.

The current speed datum can then be calculated to show the estimated available power, where the higher the current speed, the greater the power produced. The location where the power meeting is the highest compared to other places. The area is then determined as the potential location. A high enough power density potential will occur when tide conditions lead from high to low tide and on the lowest tide condition. Whereas for the small amount of power density potential only appears when the tide is highest and when the tide is lowest.

3.4. Spatial Result of Power Density (watts) with Tidal Condition

From Figure 14 and 15 we can see the result of the current speed from the numerical simulation that has been converted into a spatial power density condition at each tidal condition. We can see that in the Northern area of the Karimun Jawa Island has the highest distribution of potential Electricity Power distributed with the average Electricity Power is 450 Watt.

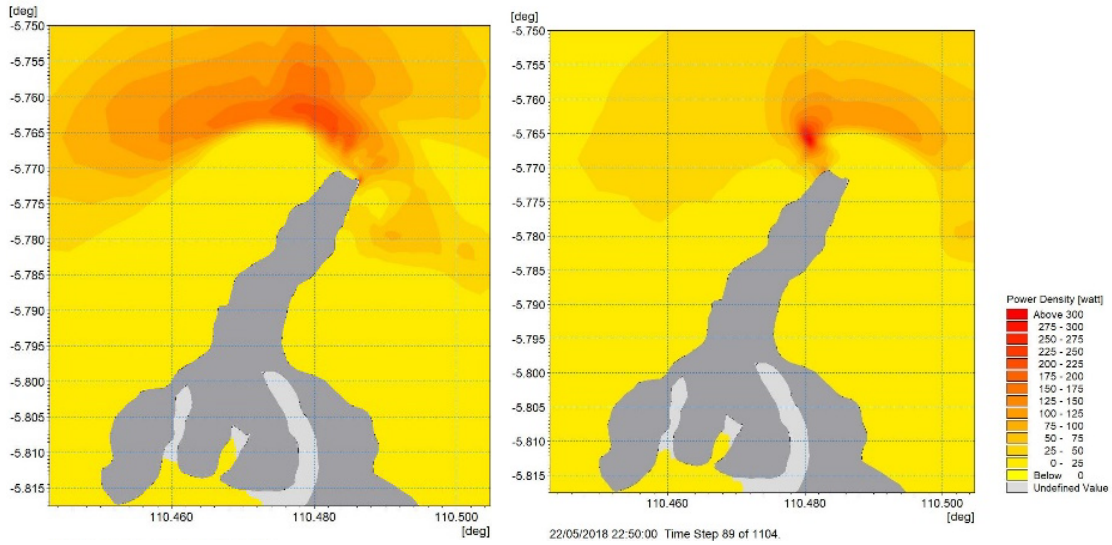


Figure 14. Spatial electrical power density result in low tides to high tides condition (left), high to low tides condition (right) in the spring season

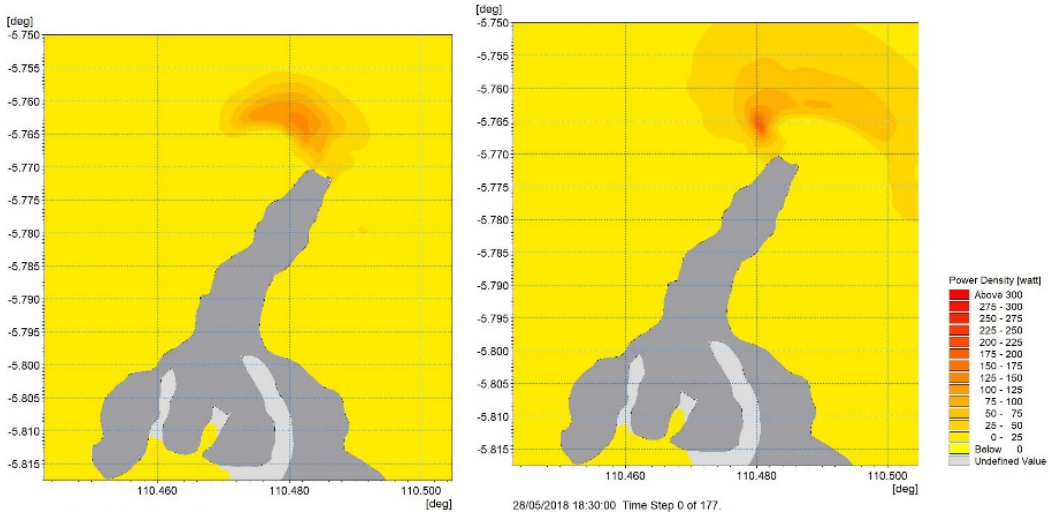


Figure 15. Spatial electrical power density result in low tides to high tides condition (left), high to low tides condition (right) in) in the neap season

4. Conclusion

potential in the West season is 565,445 Watt, and the lowest Electrical Power potential in Transition 1 season is 359.85 Watt. Based on the results of this study, it can be said that the Karimunjawa island which initially relied on PLTD for the fulfillment of its electricity powers, could have great potential to develop electricity from ocean currents, which is far more economical and environmentally friendly compared to fossil energy.

5. Acknowledgment

We would like to thanks to Ministry of Research, Technology and Higher Education which gave us funding from Fundamental Funding of Directorate of Research and Community Services, Directorate of General Research and Development, Ministry of Research, Technology and Higher Education as the agreement letter of Research Implementation assignment, with Research Number : 022 / SP2H / LT / DRPM / II / 2016 dated on 17th February 2016. Give thanks also to the students of the Oceanographic Department who involved in this research, i.e., Satrio Srijati, Albert Gunawan, Larosa Nurfikri, Muhammad Noerman and Radhitya Rega.

References

[1]. Mulyanef, Marsal, Rizky Arman, K Sopian, "Sistem Distilasi Air Laut Tenaga Surya Menggunakan Kolektor Plat Datar Dengan Tipe Kaca Penutup Miring", Teknik Mesin Universitas Bung Hatta; Padang, 2006.

[2]. Nurhayati, "Distribusi Vertikal Suhu, Salinitas dan Arus di Perairan Morotai, Maluku Utara, Oseanografi dan Limnologi di Indonesia",2006.

[3]. Departemen Kehutanan, Balai Taman Nasional Karimunjawa, "Penataan Zonasi Taman Nasional Karimunjawa Kabupaten Jepara Provinsi Jawa Tengah", 2004.

[4]. Parwito, "Tiap Hari, Listrik di Pulau Karimunjawa Hanya Nyala 6 Jam", Merdeka.com, Thursday, 20th February, 2014.

[5]. Humas Jateng, " Listrik Untuk Karimunjawa", Jatengprov.go.id, Accessed on Friday, 4th- April 2014, 3:31 pm, 2014.

[6]. Bedard. R, "Technology characterization: Ocean wave and tidal energy", Global Marine Energy Conference, April, New York, 2008.

[7]. Lopez. A, J.A. Somolinos, L.R. Nunez, and M. Santamaria, "Modeling and Simulation of Moored Devices for Ocean Currents Energy Harnessing", Journal of Maritime Research, VIII (1): 19 - 34, 2011.

[8]. Asosiasi Energi Laut Indonesia (ASELI), "Potensi Sumber Daya Energi Laut di Indonesia", ASELI, Bandung, 14 pages, 2011.

[9]. Lubis. S, "Pembangkit Listrik Tenaga Arus Laut bagi Desa Pesisir Tertinggal (Second Opinion) Ministry of Energy and Mineral Resources", 22nd April, 2014

[10]. Yuningsih. A, and A. Masduki, "Potensi Energi Arus Laut untuk Pembangkit Tenaga Listrik di Kawasan Pesisir Flores Timur, NTT", Jurnal Ilmu dan Teknologi Kelautan Tropis, 3(1): 13-25, 2011.

[11]. Siagian, Hendry Syahputra, I. B. Prasetyawan, D.H. Ismunarti, and R. B. Adhitya, "Study of the Potential of Ocean Current as an Energy Power Plant in the Larantuka Strait, East Flores, East Nusa Tenggara", Buletin Oseanografi Marina Januari 2014 Vol, 3(1): 1–8, , 2014 http://ejournal.undip.ac.id/index.php/buloma

[12]. Kementerian Energi dan Sumber Daya Mineral, "Peraturan Menteri Energi Dan Sumber Daya Mineral Republik Indonesia No. 13 Tahun 2015" Jakarta, 2015.

[13]. Sugianto, D.N and D.S. Agus, "Studi Pola Sirkulasi Arus Laut di Perairan Pantai Provinsi Sumatera Barat". Jurnal ILMU KELAUTAN, June 2007, Vol, 12 (2): 79 – 92, 2007.

[14]. Istiarto, "Model Hidrodinamika di Bidang Hidraulika Saluran Terbuka", http://istiarto.staff.ugm.ac.id. Accessed on 12 April, 2015.

[15]. Erwandi, "Sumber Energi Arus: Alternatif Pengganti BBM, Ramah Lingkungan, dan Terbarukan", Laboratorium Hidrodi-namika Indonesia, BPP Teknologi, 2006.

[16]. Priyono, "Metode Penelitian Kuantitatif", Revised Edt, Zifatama Publishing, Surabaya, ISBN : 978-602-6930-31-6. 206 pp, 2016.

[17]. Ismunarti, Dwi. Haryo, Sugianto. D.N, Ismanto. A., "Kajian Karakteristik Arus Laut Di Kepulauan Karimunjawa Jepara", Prosiding Seminar Nasional Hasil-Hasil Penelitian Perikanan Dan Kelautan Ke-Vi Fakultas Perikanan Dan Ilmu Kelautan – Pusat Kajian Mitigasi Bencana Dan Rehabilitasi Pesisir, Undip, 254 263, 2017 .

[18]. Pond, S. and G.L. Pickard, "Introductory Dynamic Oceanography Second Edition", Pergamon Press, Oxford, 329 p, 1983.

[19]. Hadi, S. dan I.M. Radjawane, "Arus Laut", Institut Teknologi Bandung, Bandung, 88 pages, 2009.

[20]. Kowalik and Murty, "Numerical Modeling of Ocean Dynamics", World Scientific, Volume 282, 10 January 1995, p. 407, 1993, https://doi.org/10.1017/S0022112095240187

[21]. Hagerman, G., B. Polagye, R. Bedard, and M. Previsic., "Methodology for Estimating Tidal Current Energy Resources and Power Production by Tidal In-Stream Energy Conversion (TISEC) Devices", EPRI, Virginia, 52 p, 2006.

[22]. Moreno, N., R. Sallent, A. Espi, D. Bao, and Y. Teillet., "Ocean Current's Energy: How to Produce Electrical Energy Thanks To The Marine Currents?", University of Gävle, Gävle, 23 p, 2008.

[23]. Fraenkel, P.L., "Marine Current Turbines: Pioneering The Development of Marine Kinetic Energy Converters", Proc. Instn. Mech. Engrs., 221 (Part A) :159-169, 2007.
[24]. Fraenkel, P.L, "Power From Marine Currents", Proc. Instn. Mech. Engrs., 216 (Part A):1-14, 2001.
[25]. Thomson E.R. dan Emery J.W., Data Analysis Methods in Physical Oceanography ,1st ed. Pergamon. England, 1997.
[26]. Gorlov, A.M., "Turbines with a twist", In: Kitzinger U and Frankel EG (eds) Macro Engineering and the Earth: World Projects for the Year 2000 and Beyond, pp, 1} 36, Chichester: Horwood Publishing, 1998.
[27]. Giorgi, S. and J.V. Ringwood, "Can Tidal Current Energy Provide Base Load?", Journal of Energies, 6: 2840-2858, 2013.
[28]. EVANS, G. P., "A Framework for Marine and Estuarine Model Specification in the UK", Foundation for Water Research, Report No FR0374, 1993.
[29]. Trujilo, A.P., and H.V. Thurman, "Essentials of Oceanography 10th Edition", Prentice-Hall, Boston, 551pp, 2011.
[30]. Cahyana, C, "Model Hidrodinamika Laut", BATAN, Jakarta, Buletin LIMBAH 9 (2) : 17-18 pp, 2005.
[31]. Djunarsjah, E., dan Poerbandono, "Survei Hidrografi", PT. Refika Aditama, Bandung, 2005.
[32]. Ongkosono, O.S.R, and Suyarso, "Pasang Surut", Pusat Pengembangan Oseanologi Lembaga Ilmu Pengetahuan Indonesia, Jakarta, 257 pp, 1989.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 133-137 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Big Data Analytics Using Deep LSTM Networks: A Case Study for Weather Prediction

Shweta Mittal[*,1], Om Prakash Sangwan[2]

[1]*Reseach Scholar*, *Department of Computer Science and Engineering, Guru Jambeshwar University of Science & Technology, Hisar, Haryana, India,125001*

[2]*Professor*, *Department of Computer Science and Engineering, Guru Jambeshwar University of Science & Technology, Hisar, Haryana, India, 125001*

ARTICLE INFO

Article history:
Received: 23 November, 2019
Accepted: 31 January, 2020
Online: 09 March, 2020

Keywords:
Big Data
Spark
Weather Prediction
LSTM
RNN

ABSTRACT

Recurrent Neural Networks has been widely used by researchers in the domain of weather prediction. Weather Prediction is forecasting the atmosphere for the future. In this proposed paper, Deep LSTM networks has been implemented which is the variant of RNNs having additional memory block and gates making them capable of remembering long term dependencies. Fifteen years hourly meteorological data of Brazil weather stations for the period of 2000-2016 collected from Kaggle.com has been analyzed for 1 hour and 24-hour time lag using Keras libraries on Spark framework. Hidden layers of the network have been increased up to three to examine its impact on accuracy of the network and it was found that network with 2 hidden layers provides good accuracy in lesser learning runtime. From the experimental results, it is also concluded that Adam optimizer provides best results when compared with SGD and RMSProp optimizer.

1. Introduction

This paper is an extension of work based on Machine Learning Techniques for Big Data Analytics originally presented in 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence) [1]. Weather prediction is one of the vital applications of technology to forecast the atmosphere for the future and has its wide application in number of domains such as forecasting air flights delay, satellites launching, crop production, natural calamities etc. Weather prediction can be classified into 4 categories depending on the time of forecast : a) Very Short Term Prediction: forecast for few hours in advance b) Short Term: forecast for days in advance c) Medium Term: forecast for weeks in advance d) Long Term: for months-years in advance. In the early days, data collected from several weather stations were fed into some numerical models which made the use of statistics to predict the future. While with the growing popularity and tremendous applications of Machine Learning Techniques, availability of high speed processors, new improved tools and techniques, researchers are now experimenting with ML techniques in meteorological domain too.

Neural Networks is one of the most popular Machine Learning techniques but it lacked its popularity in the late 90's due to gradient descent problem and limited processing power of the computers leading to very high learning time. As the technology progressed further, various techniques like parallel processing, High Performance Computing clusters, GPUs etc. have been developed which when applied in NN results in tremendous improvement of their performance.

RNN are the recurrent neural networks with the feedback connection but they might get trapped in Vanishing Gradient Problem. LSTM (Long Short Term Memory neural networks, introduced by Hochreiter and Schmidhuber in 1997) are gated RNNs which overcomes the above-mentioned issue by creating the gradients that neither explode nor vanish with time. It is composed of 1 memory unit and 3 gates namely input, output and forget gate which uses sigmoid activation function to decide which information they need to input, output and forget. LSTM has self-loops with conditioned weights controlled by forget gate to generate paths where gradient can flow for greater time [2]. It has also an advantage of learning long term dependencies, thereby widely used in text sequencing domain. Architecture of LSTM

* Shweta Mittal, Av.Email: shwetamittal019@gmail.com

networks has been described in Figure 1. LSTMs are trained using BPTT i.e. Back Propagation through Time, thus overcomes vanishing gradient problem but it might get stuck in local optima.

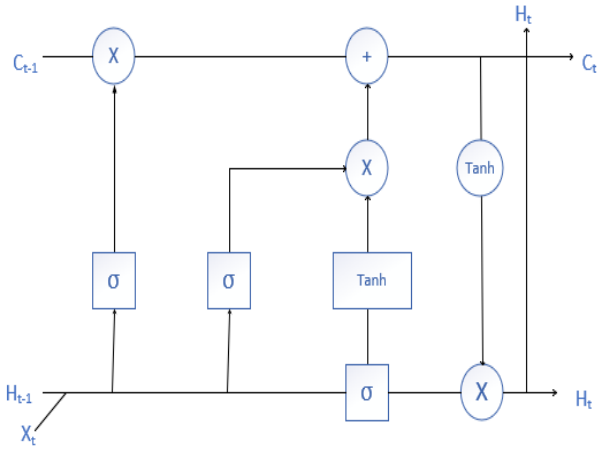


Figure 1: Architecture of LSTM: Image Source [3]

2. Literature Review

In this section, a detailed review has been conducted on work done by various researchers for application of NN in the field of weather prediction. A study on various Data Mining algorithms popularly used in the field of weather prediction has been conducted and several challenges that may occur while applying traditional data mining techniques were highlighted [4]. In [5], the authors analyzed 30 years of Hong Kong metrological data using 5 layered Deep NN with 4 layered Stacked Auto Encoders to learn raw data features with SVR for prediction. NMSE (Normalized Mean Square Error), R Square and DS (Directional Symmetry) metrics have been used for evaluation purpose. In [6], the reviewers analyzed weather data collected from www.weatherunderground.com which was first normalized and then fed to 3 layered BPNN model to obtain non-linear relationships.

IGRA dataset collected across 60 stations was analyzed by authors to design data driven kernel and to provide an efficient inference procedure [7]. Best validation results were obtained using 2 stacked RBM with hidden neurons 50 and 150. In [8], the authors used CNN for the first time to predict monthly rainfall of Australia and the results proved that CNN performs better in comparison to prediction model used by Bureau of Meteorology and MLP for months with high annual average rainfall. Researchers also proposed CFPS-HGA (Climate features and neural network parameters selection-based hybrid genetic algorithm) for setting neural network parameters and the proposed algorithm performed better [9].

Meteorological data collected from BKMG (for the period 1973-2009) and ENSO data from NOAA was analyzed by researchers using RNN, CRBM and Convolutional Networks [10]. In [11], the authors implemented NARX NN (Non Linear Auto regressive with exogenous variables), a type of recurrent NN for predicting multiple weather attributes and the results concluded that proposed algorithm performed better than CBR (Case Based Reasoning). In [12], researchers proposed a new layer i.e. Dynamic Convolutional layer which used filter in addition to feature maps from the previous layer to predict the sequence of rain radar images using Caffe library and Stochastic Gradient Decent optimizer.

In [13], the researchers compared the performance of deep neural NN with shallow NN for short term wind prediction. NN parameters were reduced using PMI (Partial Mutual Information) based IVS algorithm and the results proved that 3 layered stacked auto encoders performed better. In [14], the authors proposed Deep Multilayer Perceptron using Cascade learning and implanted it on H1-B Visa applications record and cardiac dataset and the results proved that proposed model has less run time and greater accuracy.

In [15], the authors proposed 2 layers stacked LSTM model (each layer has independent LSTM model for every location) for weather data obtained from Weather Underground website from 2007 to 2014. In [16], the authors compared the performance of ARIMA model and LSTM (weather variables are combined) on weather data collected from Weather Underground for the period 2012 to 2016 and it was concluded that LSTM performs better than ARIMA.

In [17], the authors proposed deep LSTM model for wind prediction on the data from Manchester Wind Farm for 2010-2011 and the proposed algorithm performed better than BPNN and SVM. Best results were obtained from the model consisting of 3 hidden layers constituting 300, 500 and 200 neurons. In [18], 15 years of hourly meteorological data for 9 cities in Morocco was analyzed by authors using multi-stacked LSTM to forecast 24 and 72 hours weather data. Missing values were eliminated using forward filling and RMSProp optimizer was used.

3. Experimental Study

As per the survey done in Section 2, it can be concluded that LSTM is one of the most popular and successfully used technique in the domain of weather prediction. Thus, in this experimental study, multi-stacked LSTM model has been implemented on Spark platform to determine the best architecture for the task of weather prediction. For the analysis purpose, dataset of 2 GB from INMET (National Meteorological Institute - Brazil) has been used which consists of hourly weather data from 122 weather stations of various south-eastern states of Brazil i.e. Rio de Janeiro, Sao Paulo, Minas Gerais and Espirito Santo from 2000 to 2016 [19]. Brazil is the largest country in the South America and its temperature rarely fall below 20 degrees. In the proposed experiment, a computer with 1TB memory, 16 GB RAM and 1.8 GHz processor has been used. Dataset includes 31 features and around 9.7 million (9,779,168) rows. All the dataset is first loaded into the Spark dataframe and unnecessary columns i.e. wsnm, wsid, elvt, lat, lon, inme, city, prov, mdct, date, yr, mo, da, hr, tmax, tmin have been deleted from our schema (Only 15 attributes have been preserved as mentioned in Table 1).

Data pre-processing is the pre-requisite of any analytics process. All the null values have been replaced with zero and the

results are cited in Table 2. Several other ways to deal with null values are as follows: a) Replace nulls with zero in case of numerical attributes b) Replace with mean/median/mode values c) Use algorithms that support missing values d) Remove the rows having null values.

Table 1: Attributes for weather prediction

Attribute	Description
Prcp	Amount of precipitation in millimeters (last hour)
Stp	Air pressure for the hour in hPa to tenths (instant)
Smax	Maximum air pressure for the last hour in hPa to tenths
Smin	Minimum air pressure for the last hour in hPa to tenths
Gbrd	Solar radiation KJ/m2
Temp	Instant Air Temperature (Celsius degrees)
Dewp	Instant Dew Point (Celsius degrees)
Dmax	Maximum Dew Point (Celsius degrees)
Dmin	Minimum Dew Point Temperature (Celsius degrees)
Hmdy	Relative Humidity of Air (%
Hmax	Maximum Relative Air Humidity (%)
Hmin	Minimum Relative Air Humidity (%)
Wdsp	Instant Wind Speed (meters per second)
Wdct	Wind direction in radius degrees (0-360)
Gust	Wind Gust Intensity (meters per second)

Input to the proposed LSTM network is prcp, stp, smax, smin, gbrd, dewp, dmax, dmin, hmdy, hmax, hmin, wdsp, wdct, gust, temp at $(t-1)^{th}$ hour and the output is temperature at $(t)^{th}$ hour and at $(t+24)^{th}$ hour. Architecture of the proposed network has been described in Figure 2. Input attributes are normalized between 0 and 1 via pyspark.ml.MinMaxScaler and training and validation data has been split in the ratio 80:20.

4. Experimental Results

Let I, H and O be the number of neurons in input layer, hidden layer and output layer respectively. Models are trained using Keras module (with Tensorflow backend) which has dense library for implementing LSTM networks. Weights to be optimized in the LSTM networks are as follows:

W1=4*I*H; U1=4*H*H;
B1=4*H; W2=H*O;
B2=O .....(1)

For H=5 (5 units in LSTM with 1 hidden layer), I=15 (15 input attribute) and O=1 (1 output attribute), W1= 4*5*15=300, U1=4*5*5=100, B1=4*5=20, W2= 5*1=5, B2=1. Thus, there are total of 426 parameters which needs to be optimized.

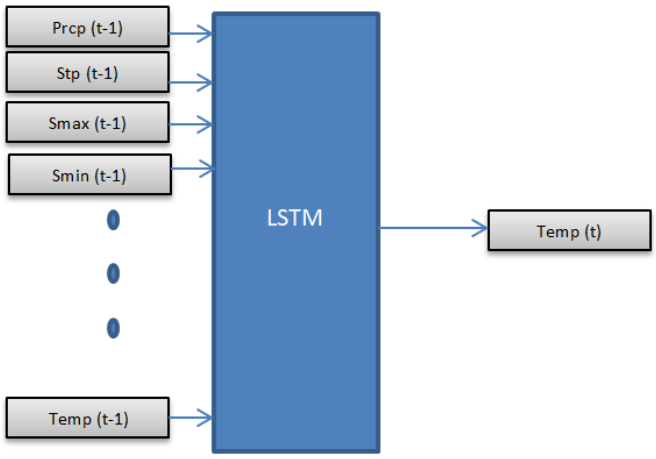


Figure 2: Architecture of LSTM network

Table 2: Number of nulls

Attributes	Number of null values
Prcp	8371184
Stp	0
Smax	0
Smin	0
Gbrd	4108820
Dewp	475
Dmax	310
Dmin	807
hmdy	0
Hmax	12
Hmin	44
Wdsp	925561
Wdct	0
Gust	316474
Temp	31

Two separate networks have been implemented, one provides output temperature at t^{th} hour and the other network provides output at $(t+24)^{th}$ hour. In the proposed study, the number of hidden neurons has been varied from 10 to 30 and the number of hidden layers has been increased up to 3. Results of both the case studies have been mentioned in Table 3. Performance of Adam (Adaptive Moment Estimation), SGD (Stochastic Gradient Decent) and RMSProp (Root Mean Square Prop) optimizers has also been compared for 2 hidden layers comprising 10 and 20 neurons respectively for 24 hours' time lag and results are described in Table 4. From the results, it can be concluded that Adam optimizer performs best whereas SGD performs worst. Thus, Adam optimizer has been used in all the implemented case studies. Mean Absolute Error (MAE) metrics has been used for the evaluation purpose. Run time of the network (in floating point number) has also been analyzed to determine the complexity of the network.

Table 3: MAE of LSTM networks with 1 hour and 24-hour time lag

S. No	Number of hidden layers	Number of hidden units	Iterations	1 hour time lag		24 hour time lag	
				MAE	Runtime	MAE	Runtime
1	1	10	100	0.015151	28041.47	0.041649	30256.38
2	1	20	100	0.015462	29292.39	0.041582	29889.34
3	1	30	100	0.015365	30854.15	0.042124	28289.06
4	2	10,10	100	0.015088	48892.43	0.040463	49878.65
5	2	20,20	100	0.015077	49898.42	0.040839	46737.57
6	2	30,30	100	0.015259	50805.43	0.040060	46829.74
7	2	10,20	100	0.015124	42302.25	0.040020	52487.23
8	3	10,10,10	100	0.015269	124915.8	0.040546	101543.68

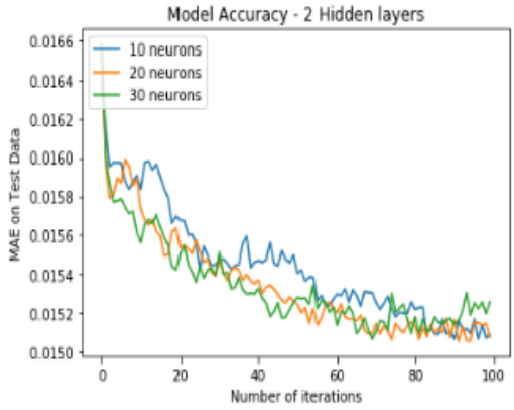


Figure 3: MAE of LSTM with 2 hidden layers for 1-hour time lag

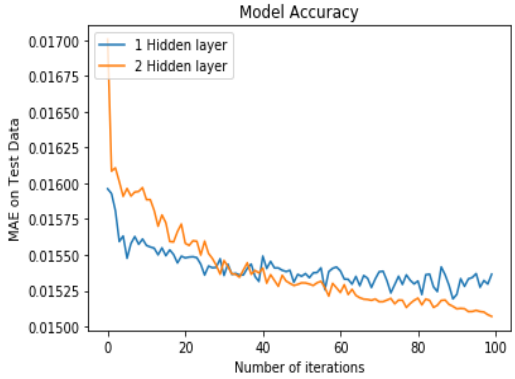


Figure 4: MAE with 1 and 2 hidden layers, each having 30 neurons each for 1 hour time lag

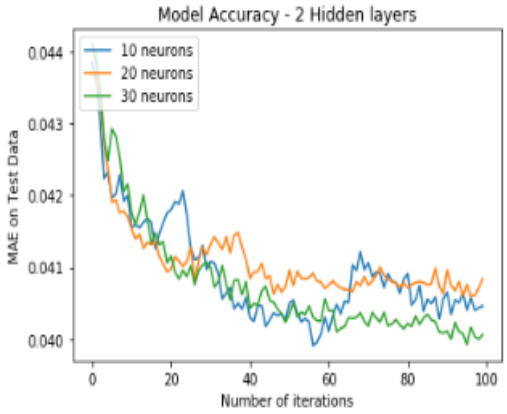


Figure 5: MAE of LSTM with 2 hidden layers for 24-hour time lag



Figure 6: MAE with 1 and 2 hidden layer, each having 30 neurons each for 24 hour time lag

Table 4: comparison of MAE for various optimizers for 24-hour time lag

Optimizers/Metric	MAE	Accuracy	Runtime
Adam	**0.04002**	6.64686E-06	52487.2306
Rmsprop	0.05972	5.62427E-06	51896.5460
SGD	0.04142	6.64686E-06	40895.8979

Figure 3 and Figure 5 illustrates the comparison of MAE with 2 hidden layers comprising 10, 20 and 30 neurons each for 1 hour time lag and 24 hour time lag respectively over 100 iterations. For 1 hour time lag, network with 2 hidden layers comprising 20 neurons each provides best results (with the lowest MAE value of 0.015077 as stated in Figure 3) whereas for 24 hour time lag,

network with 2 hidden layers comprising 10 and 20 neurons respectively provides best results with the MAE value of 0.040020.

Figure 4 shows the comparison for networks of 1 hidden layer having 30 neurons with the network of 2 hidden layers comprising 30 neurons each for 1 hour time lag and it can be inferred that network with 2 hidden layers has low MAE value as compared to network with 1 hidden layer. Similarly, Figure 6 shows comparison for networks of 1 hidden layer having 30 neurons with the network of 2 hidden layers comprising 30 neurons each for 24 hour time lag and it can be inferred that network with 2 hidden layers has low MAE i.e. 0.040060 value as compared to network with 1 hidden layer i.e. MAE value of 0.042124.

From the case studies implemented, it can also be observed that MAE for 1 hour time lag is comparatively low, thus gives more accurate results as compared to network with 24 hour time lag i.e. network with short term predictions provides higher accuracy as compared to network with medium term predictions.. More combinations of the hidden neurons and layers can be explored in the near future to obtain better results.

5. Conclusion and Future Work

In this paper, Deep LSTM network has been implemented on 15 years meteorological data. Number of hidden layers has been increased up to 3 and it can be concluded that increasing the number of hidden neurons in the layer does not necessarily increases the accuracy of the model as network might start learning the linear relationship between the attributes. Also, increase in the number of layers in the model increases the run time of network by marginal amount, thus increasing the complexity of the network.

There are number of combinations possible for choosing the adequate number of hidden neurons and hidden layers in the network which can provide the minimum MAE. Manually testing all such combinations will be a quite tedious task. Thus, process of selection of number of hidden layers and hidden neurons needs to be automated.

Also, the algorithm takes quite longer time to run on single CPU. Thus, it becomes hard to increase the number of hidden layers and neurons beyond a significant point. Parallel processing of neural network might help us in reducing the time complexity of the network. Time complexity can further be reduced by employing more number of processors to do a task or by using the high speed GPUs. In future, it would be interesting to study the reduction in time complexity by implementing the above algorithms on increased number of processors and on GPUs.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] S. Mittal, O.P. Sangwan,"Big Data Analytics using Machine Learning Techniques" in 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India. http://dx.doi.org/ 10.1109/CONFLUENCE.2019.8776614.

[2] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, 2016.

[3] https://en.wikipedia.org/wiki/Long_short-term_memory.

[4] S. Mittal, O.P. Sangwan," Big Data Analytics Using Data Mining Techniques: A Survey" in book: Advanced Informatics for Computing Research,264-273, 2018 Second International Conference, ICAICR, Shimla, India. http://dx.doi.org/ 10.1007/978-981-13-3140-4_24.

[5] J. Liu, Y. Hu, J. Jia You, P. Wai Chan,"Deep Neural Network Based Feature Representation for Weather Forecasting" in 2014 International Conference on Artificial Intelligence (ICAI), Las Vegas, Nevada, USA.

[6] Ch. J. Devi, B.S. Prasad Reddy, K.Vagdhan Kumar, B.Musala Reddy, N. Raja Nayak,"ANN Approach for Weather Prediction using Back Propagation" International Journal of Engineering Trends and Technology, 3(1), 20-23, 2012.

[7] A. Grover, A. Kapoor, E. Horvitz,"A Deep Hybrid Model for Weather Forecasting" in 2015 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD, 379-386, Sydney, NSW, Australia. http://dx.doi.org/10.1145/2783258.2783275.

[8] A. Hiader, B. Verma,"Monthly Rainfall Forecasting using One-Dimensional Deep Convolutional Neural Network," IEEE Access, 6, 69053 – 69063, 2018. http://dx.doi.org/10.1109/ACCESS.2018.2880044.

[9] A. Hiader, B. Verma,"A novel approach for optimizing climate features and network parameters in rainfall forecasting" Soft Computing, Springer, 22(24), 2018. https://doi.org/10.1007/s00500-017-2756-7.

[10] A. Galih Salman, B. Kanigoro, Y. Heryadi,"Weather forecasting using deep learning techniques" in 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia. http://dx.doi.org/10.1109/ICACSIS.2015.7415154.

[11] S. Biswas, N. Sinha, B. Purkayastha, L. Marbaniang,"Weather prediction by recurrent neural network dynamics" Int. J. Intelligent Engineering Informatics, 2(2/3), 166-180, 2014. http://dx.doi.org/10.1504/IJIEI.2014.066208.

[12] B. Klein, L. Wolf, Y. Afek,"A Dynamic Convolutional Layer for Short Range Weather Prediction" in 2015 IEEE, Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA. http://dx.doi.org/ 10.1109/CVPR.2015.7299117

[13] M. Dalto, J. Matusko, M. Vasak,"Deep neural networks for ultra-short-term wind Forecasting" in 2015 IEEE, International Conference on Industrial Technology (ICIT), Seville, Spain. http://dx.doi.org/ 10.1109/ICIT.2015.7125335.

[14] A. Gupta ,H. Kumar Thakur , R. Shrivastava, P. Kumar, S. Nag, "Big Data Analysis Framework Using Apache Spark and Deep Learning" in 2017 IEEE, International Conference on Data Mining Workshops (ICDMW), New Orleans, LA, USA. http://dx.doi.org/ 10.1109/ICDMW.2017.9.

[15] Z. Karevan, J. Suykens,"Spatio-temporal Stacked LSTM for Temperature Prediction in Weather Forecasting" Cornell University, 2018.

[16] A. Galih Salman, Y. Heryadi, E. Abdurahman, W. Suparta,"Weather Forecasting Using Merged Long Short-Term Memory Model (LSTM) and Autoregressive Integrated Moving Average (ARIMA) Model" Journal of Computer Science, 14(7), 930-938, 2018. https://doi.org/10.3844/jcssp.2018.930.938

[17] Q. Xiaoyun, K. Xiaoning, Z. Chao, J. Shuai, M. Xiuda,"Short-Term Prediction of Wind Power Based on Deep Long Short-Term Memory" in 2016 IEEE, PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Xi'an, China.http://dx.doi.org/10.1109/APPEEC.2016.7779672.

[18] M. Akram Zaytar, C. El Amrani,"Sequence to Sequence Weather Forecasting with Long Short-Term Memory Recurrent Neural Networks" International Journal of Computer Applications, 143(11), 2016. http://dx.doi.org/10.5120/ijca2016910497.

[19] https://www.kaggle.com/PROPPG-PPG/hourly-weather-surface-brazil-southeast-region.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 134-139 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Spiral Curve for Revocable Touchless Fingerprint Template Securisation

Tahirou Djara[1,2,*], Boris Sourou Zannou[1,2], Antoine Vianou[1,2]

[1]*Laboratoire d'Electronique de Télécommunication et d'Informatique Appliquée (LETIA/EPAC), Université d'Abomey-Calavi (UAC), Cotonou, Bénin*

[2]*Institut d'Innovation Technologique (IITECH), Abomey-Calavi, Bénin*

ARTICLE INFO

Article history:
Received: 03 August, 2019
Accepted: 07 November, 2019
Online: 25 November, 2019

Keywords:
Revocability
Secrecy
robustness

ABSTRACT

Fingerprint data is really protected by cancelable fingerprint template because it can be revoked when compromise and a new one can be reissued. We develop a touchless cancelable fingerprint template whose algorithm was published in our previous work. We implement here the algorithm and conducted several tests on several databases to confirm the stability of the model. To justify how specific keys are used, we used the Kolmogorv-Smirnov test (K-S) and the distribution histogram of legitimate / impostor scores. We compared two systems in which users register. This is the reason for the average value of K-S (0.7812) and similarity assessment (2.4703). These results improve sufficiently (6.1636 and 0.9934 for the successive separability and the K-S test) during the evaluation of the user keys through the second device. We have tested the diversity of curves that we generate. Our proposed non-contact revocable fingerprint model has demonstrated robustness against the security challenges that fingerprint authentication systems are exposed to. We evaluated it on our own database. The requirements of revocability, diversity and security are achieved with very good performance as evidenced by the FAR (False Acceptance rate) obtained on our database (0.0015).

1. Introduction

Biometry is the identification from human characteristic and traits. The most used feature in human authentication is the fingerprint. It performs well and is unique. Minutiae are the most used representation. Nevertheless, several researches have proved that one could reconstruct the original fingerprints from some characteristics. There are problems of correspondence and paring. It should be noted that during the matching phase, the model is exposed to several intrinsic security issues including the threat of privacy, attack record multiplicity(ARM) which are the most common. ARM remains the most formidable and continues to be the focus of intensive research. Therefore the issue of securing authentication with fingerprints comes out. We design a new model for data protection enhancement protection. Our vision is therefore to deal with them by taking into the paradigms of diversity, revocability, precision and invertibility.

In [1], Author identified three major classes of fingerprint-based protection models: feature transformation, biometric cryptosystem, and hybrid. Each of these methods present its own limitation and advantages [2]. The principal requirement of a good protection template hasn't been achieved. Our interests focus on rigid transformations. Basically it will be a G transformation function with original characteristics by using a key b; the transformed template G(a, b) is then stored in the database. We use that function to transform the characteristic test c, then the transformed characteristics G(a, b) and G(c, b) will be compared in order to know yes or no if the user is the right one. Entity transformation models are classified into two main classes, namely vector-based approach and point-of-interest approach (especially those based on minutiae). This depends on the representation model that is adopted.

This article proposes a new approach that is difficult to reverse, especially for systems using minutiae as model. This technic provides revocability, diversity and security while improving performance. It should be noted, however, that to ensure security through the bio-cryptosystem or transformation of characteristics, the fingerprint device observes a delay in the accuracy or during the inversion. Our device for strengthening the fingerprint authentication has the advantage of avoiding leash dragging during the image acquisition step. Our vision is to implement a non-invertible transformation that meets the requirements of

* Tahirou Djara, Email: csm.djara@gmail.com

performance, diversity, revocability and security. Thus, we use information provided by the munities obtained to generate contactless curves with respect to the center of mass. Section 1 presents the introduction, then in section 2 we present the security holes in contactless biometric systems. The section 3 present our revocable and secure contactless model. Section 4 presents the experimental results to finally give the conclusion in section 5.

2. Security holes in non-contact biometric systems

Biometric systems are exposed to many security vulnerabilities [3,4]. Malicious, for example, could attack the database and recover sensitive information. In order to secure the biometric systems of the most unimaginable attacks, intensive research efforts are conducted. Many techniques have been developed for the purpose of securing biometric data.

. The basic idea of all these models is to generate revocable templates. By revocability it should be understood that in comparison to passwords that could be modified if they are damaged, the user model stored in the database could also be generated in a new (different from the previous) using the same biometric information. Diversity, security and performance are the characteristics that an ideal model of protection must have for security models. Diversity means that it is necessary to ensure that there is no correspondence between the compromised models and those newly generated in the case where a new model is generated to replace an old one Security is the impossibility for a malicious person to recover the original data of a user who was used to build his model. Performance means keeping the authentication capacity of the biometric system intact, which should not be affected by data protection. In a biometric system, instead of storing biometric templates directly, the transformed templates are stored in the database for subsequent authentication. In this technique, the biometric characteristics are transformed into another domain [6] and only the transform (signature) is stored in such a way that the details of the original biometric data cannot be revealed to a malicious person. he gets the biometric pattern. The biometric characteristic elements are transformed into another model G (a; b) which will be stored in a database for future authentication, the biometric information β of the user is transformed into G (c, b) to allow a comparison with G (a, b) to decide whether it is a malicious or an authentic one. The attacks essentially originate in the security vulnerabilities that can essentially be counted:

- Presentation attacks: fingerprint is presented at the inputs after having reproduce it;
- hacking and using data from fingerprint after bypassing Sensor
- The usurped characteristics are substituted for the originals.
- Tripatouillage in the correspondence module to use the false features

In [1], Author have cited security loopholes through a fish skeleton.

3. Literature review

Given the security weaknesses listed above, it is necessary to look for sustainable solutions. It is within this framework that solutions approaches are born. We can mention revocable biometrics without alignment and biometrics based on pre-registration. These solutions are based on the local structures of the minutiae or the taking of the munities in pairs or in a triangle because they do not vary with the rigid transformation. In the next paragraph we will go through the revocable models without alignment. In [5], Authors have proposed one of the most recent models by inserting two key factors. Their model focused on a self-alignement local structure based on textures as features. Their research was much more focused on the many possible attacks rather than dealing with algorithms managing the loss of models or two keys. Another recent technique in [5] evaluates the orientation of the munitie in the surrounding area at each reference munitie. All this is possible thanks to two functions allowing to evaluate the number of transformations around each munitie of reference. The performance of this technique drops when it comes to poor quality images.

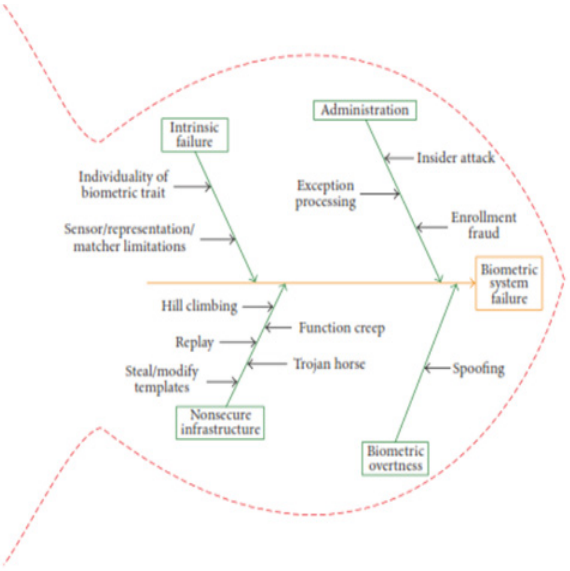


Figure 1: vulnerabilities in the form of fishbone [1]

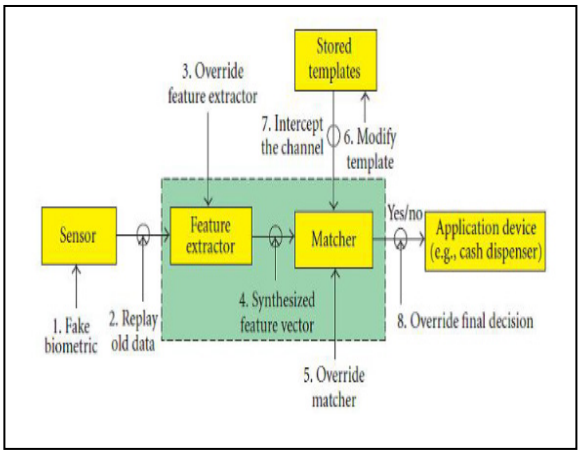


Figure 2: Attacks on Biometric System

In [6], author proposed a set of hash functions based on the position of the munities. These transformation function based on the position of the munitie, takes into account the hazardous mobility points munitie due to the sensor recording. This technique is well revocable and performs well with high complexity. The vicinity of the munities is a technique based on the dynamic random projection to secure the extracted features. It was developed by authors in [7]. The random projection matrix is dynamically assembled. The feature vector conditions the choice

of projection vectors. The test results on some databases are still insufficient to prelude because this approach would be applyed to other databases. In [8], authors used a method based on the relative position between each reference munitiae. Another munitiae based template is built in a polar coordinate system whose faults reside in poor quality performance. In [9], authors developed an approach based on a characteristic vector based on a triangle. The template is formed by quantization and binarization. A descriptor based on a string of characters, allows to set up plates in the form of binary codes. This technique is called the multiline code (MLC) developed by authors in [10] and is based on the decomposition of neighborhood munities. Similarly, models referencing mapping in an infinite approach emerged with authors in [10]. They introduced the Hamming method based on graphs generating binary formats that can be revoked. Other not less powerful techniques such as reduced circular convolution [11], the partial transformation of Hadamard [12].

Despite of plethora of methods, fingerprint authentication systems have security vulnerabilities and are still vulnerable to ARMs (Multi-Object Attack) [1].

4. Touchless secured revocable template based on center of mass

Because of weaknesses of existant template we propose a spiral curve for securing contactless fingerprint template that use Zernike moment [13], Hausdorff distance modified [14, 15, 16] and the geometric moment [17]. In this section we will present our Touchless fingerprint revocable template based on center of mass. To strengthen fingerprint authentication, We plot the contactless curve using images obtained from our contactless sensors. Later we attach them to those previously stored. then we calculate the corresponding score to finally make a decision to accept or refuse.

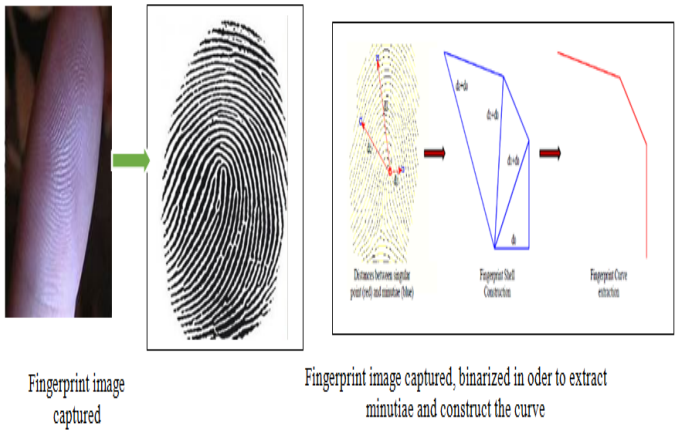


Figure 3: Building of the spiral curve

Our system is monomodal and is based on fingerprints. The rigid transformation [17] without modifying the extracted characteristics, protects the model stored there. Rigid transformation is a combination of translation and rotation expressed by:

$$\begin{pmatrix} u'-u_0 \\ v'-v_0 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\delta & -\sin\delta & 0 \\ \sin\delta & \cos\delta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & b_x \\ 0 & 1 & b_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u-u_0 \\ v-v_0 \\ 0 \end{pmatrix} \quad (1)$$

Where

$$\begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_0 & \alpha_1 & \alpha_2 \\ \beta_0 & \beta_1 & \beta_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad (2)$$

with

$$\alpha_0 = \cos\delta \quad \alpha_1 = -\sin\delta$$

$$\alpha_2 = (1-\cos\delta)u_0 + v_0\sin\delta + b_u\cos\delta - b_v\sin\delta \quad (3)$$

$$\beta_0 = \sin\delta \quad \beta_1 = -\cos\delta$$

$$\beta_2 = (1-\cos\delta)v_0 + u_0\sin\delta + b_u\sin\delta - b_v\cos\delta \quad (4)$$

Noted that $F_0\begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$ is the center of rotation, δ is the angle of rotation, $\begin{pmatrix} b_u \\ b_v \end{pmatrix}$ are the coordinate of translation vector and F' is the transformed point of F.

In oder to eliminate all the above danger, we design an new method which is implement in [3].

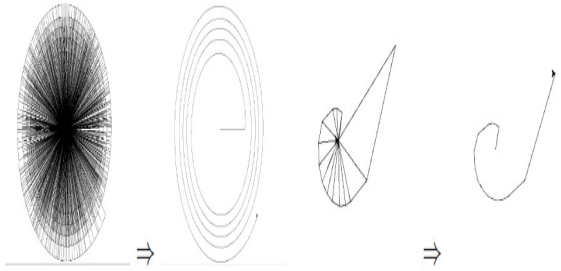


Figure 4: Different curves obtained according to the quality of the images and the number of minutiae present on these images

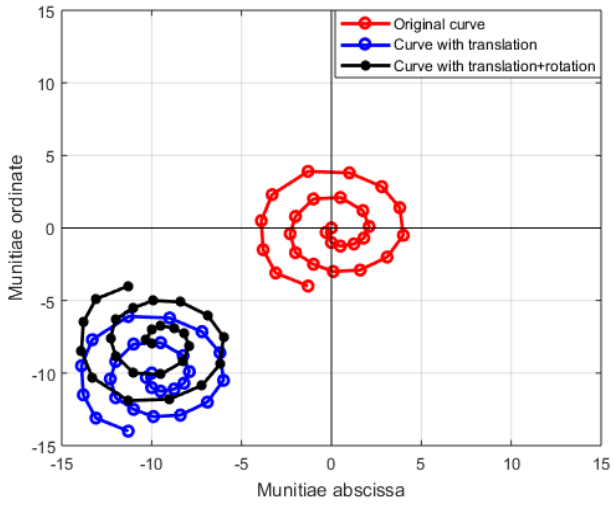


Figure 5: Construction of the secured touchless template

The requirements of revocability, diversity and security have been achieved with our new method. In reality the protection of a compromised model can go through the modification of C_0. The mismatch of the most recent curve with the old one provides diversity (see Figure 4). In reality, it is impossible to find the

munities to walk a contactless curve but it is nevertheless necessary to look at the most extreme case. In this case, even if the thug gets the curve and knows how to leave the curve to recover the munities. In reality he could only recover distances that have infinite positioning. There exists only 360 positioning and 360n possible combinations with n corresponding to the number of munities. We therefore find that in the most extreme case our model remains stable and robust despite the complexity of the attacks perpetrated. Performance, Diversity and security will be further analyzed in the next subsection.

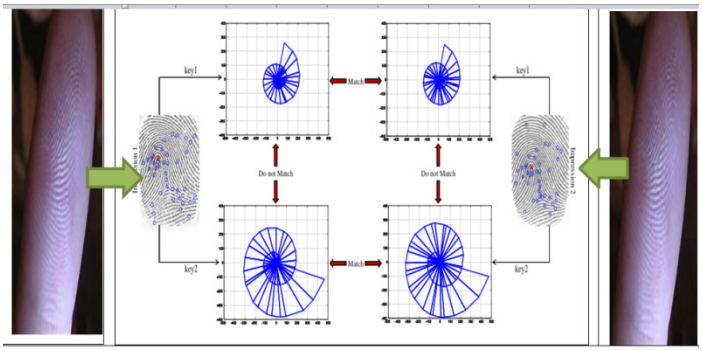


Figure 6: sample of two impressions of the same identity using two differents keys to illustrate revocability/diversity

5. Experimental results and analysis

We tested the model by using our own fingerprints database. We describe the database in Table 1. We have a fingerprint database containing 1,512 fingerprints, obtained from 378 separate fingers. We take 4 impressions from each subject. Out of 378 people, the fingerprints of the four fingerprints could be automatically captured from 1512 people. Thus only 1250 fingerprints impressions were used in experiment. In fact, due to the lack of time and people from whom we can get fingerprint images, we import some images from existant database in order to reach the number we choose. It doesn't affect the result. *We* develop our own software for the fingerprint images treatment. To appreciate how our system performs, we use false acceptance rates and false rejection rates

Due to the fact that we use two keys, we notice the amelioration of the performance of template. Our c_0 allows us to have a significant improvement of the Equal Error Rate (EER) because of the specific way of it is compute and its only one occurrence [18]. The rigid transformation applied to the curves adds a high level of randomization. Because of the fact that the key value is specific to users, the use of rigid transformation help us to have the curve unable to be rebuilt. In Figure 6, We compare FAR obtained for the proposed technique with Fingerprint Shell [17, 18] for different thresholds.

We have modified the FVC evaluation protocol to avoid the zero effort scenarios, where the adversary knows and tries to bypass the system using his own fingerprints. Impostor scores is obtain from comparison of the first curve of one single finger and the first curve of the same finger in another system that we build with reference to the identic curve. We note that they must perform 8999 attacks before succeeding one time their attacks. We compare every single print of any finger with the remaining impression of same finger. We will have at least 2799 trying before succeeding one time. In addition, we select the key randomly belonging to the interval]0, 1001[. Figure 8 describe the scenario.

Table 2: Performance comparison: existing template versus proposed template following FVC

Various Techniques	**Our collection**	**FVC2002 DB2(%)**	**FVC2002 DB1(%)**
[12]	Not tested	1.2	2.1
[15]	Not tested	3.61	7.18
[2]	Not Tested	1.01	2.03
[18]	0.001	Not tested	Not tested
Our template	0.0015	0.96	1.54

The value of FRR remains the same comparatively to fingerprint shell because of the uniqueness of the key. Let us remarks that the rightness of the curve depends on the quality of the image. More there is minutiae better is the curve, and less there is minutiae, worst is the curve (figure 9). The propose template performs well. Our model presents good results and verifies the criteria of good protection systems. Our Solution is efficient against attacks without effort or even in case of usurpation of the user key. Although, as shown in Table it presents the values EER, FMR1000 and zeroFMR obtained. Performance is only maintained in the vicinity of ERR points (4.2705% for DB1 and 1.458% for DB2). Our template are no longer accurate for thresholds that are far from ERR threshold (figure 7). In any case, we retain that (in the FMR security applications of Table 3), the systems for securing which are based on the curves are efficient only if the threshold is close to the EER point in order to limit the passage in force of the effortless attacks.

Table 3: Verification of accuracy of our template in the zero effort attack:

	FVC 2004			FVC 2005	
	EER	FMR_{100}		EER	FMR_{100}
Our touchless Secure template	4,250	26,146	Our touchless Secure template	4,250	26,146

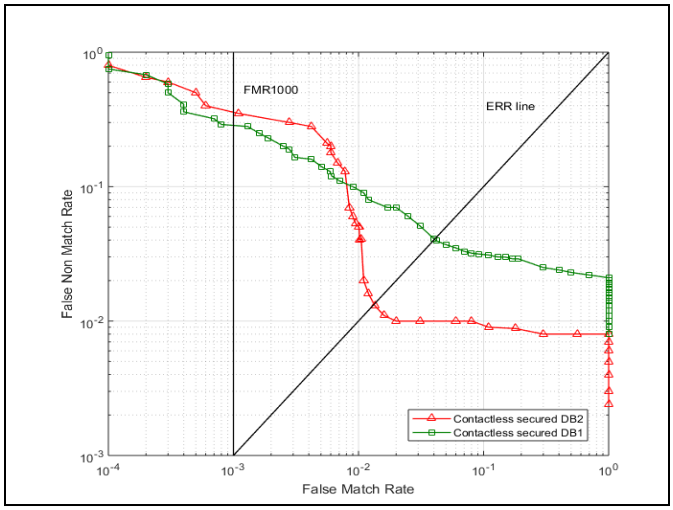


Figure 7: ROC curve in the presence of effortless attack

We remark that in high security applications based on their scheme, the systems should be operated only near the EER point to minimize the success of zero effort attacks.

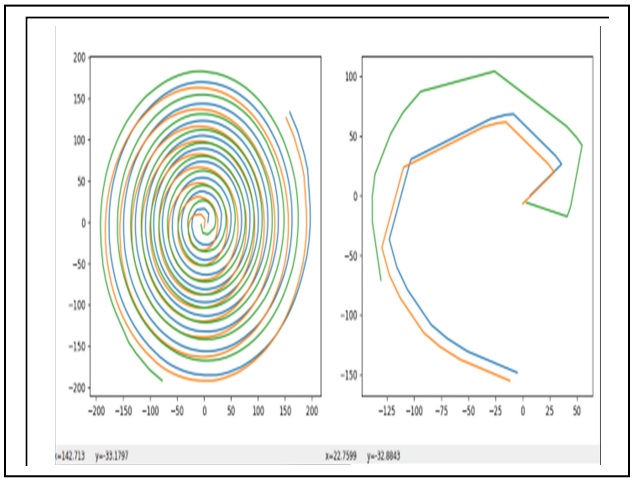


Figure 8: Sample of Template obtained from two kind of image: the first template is obtained from an image with many munitiae vs the second one is obtained from an image without enough minutiae.

Given the poor quality of some images, we observed 7 rejections for DB2 and 21 for DB1. Thus, we obtained REJDB1 = 0.75% and REJDB2 = 0.25%. With the intention of testing the opportunity to make use of C_0 particular to each user, we calculate the separability.

To justify the way in which specific keys are used, we used the kolmogorv-Smirnov test and the distribution histogram of legitimate / impostor scores. We compared two systems in which users register. In the first, the registration is done without user key, while in the second, a user key is required (the keys are drawn randomly in the interval [0; 1,5555]). We have illustrated the results in Figure 9. We can notice a form of superposition between the two distributions. This is the reason for the average value of the K-S (0.7812) and the similarity evaluation (2.4703). these results improve sufficiently (6.1636 and 0.9934 for successively the separability and the K-S test) during the evaluation of the user keys through the second device (Figure 9-b). This is proof of the improvement of discredit. In conclusion, the rate of false acceptance improves when we assigns to each user a particular key C_0. Thus, we improve the performance and accuracy of our device.

In a second step, we want to test the diversity of the curves that we generate. It will be a matter of reassuring oneself that the curve generated from one finger in one system does not coincide with another curve generated in other systems with the same finger. For this purpose, we have installed three devices that use loops. For randomly selected users in the ranges of [0,1.5555], [100,500] and [1000,2000]. The same finger is used to register in all three systems and then we look for pseudo-legitimate distribution. We conducted 6355 tests for each device, We did 6355 attempts at database DB1 and 6385 attempts for the database DB2. Figure 8 shows the results or we can clearly see that the real users and the malicious ones are well separated in the distribution of our first device. The closer they are to the distribution of the malfrats of the device 1 so the device 1 always considers the curves of the devices 2 and 3 as usurpers.

To summarize, we will say that revocability and diversity are taken into account by our device. The pseudo-distribution of the device 3 is further from the distribution of the real users of the device 1 and 2, therefore the diversity of two devices that rely on our model increases when the choice intervals of the keys are distinct.

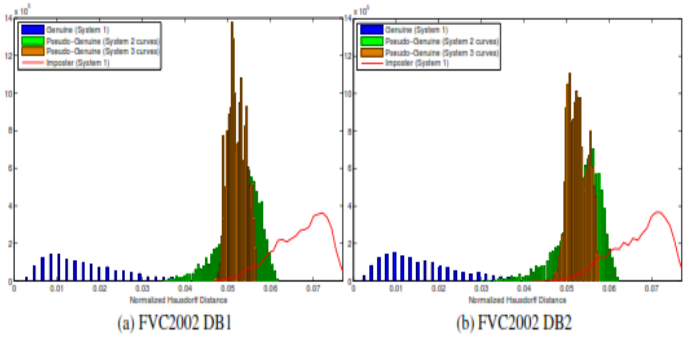


Figure 9: System Histogram 1 and pseudo-legitimate distributions using systems 2 and 3

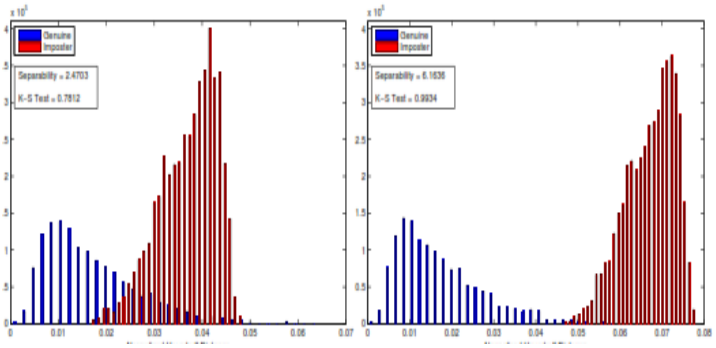


Figure 10: Histogram of Two Systems Protected by Fingerprint Shell Using FVC2002 DB1.

6. Conclusion

Our proposed contactless revocable fingerprint template has shown good robustness against the security challenges that fingerprint-based authentication systems are exposed to. It also bypasses the contagion hazards to which contact systems expose us. We evaluated it on FVC2002 DB1, DB2 and DB3 and FVC2994 DB2 and the results of the tests show that it performs very well compared to existing models in the literature review. the exisgences of revocability, diversity, security are reached with very good performances. ARM issues will be better addressed with proposals for more stable models such as exploring a combination of DFTs and dynamic projection.

Conflict of Interest

The authors declare no conflict of interest.

References

[1]. Jain, A. K., Nandakumar, K. and Nagar, A. 'Biometric Template Security', (January), pp. 1–20, 2008.

[2]. Rathgeb, C., Uhl, A., A survey on biometric cryptosystems and cancelable biometrics. EURASIP Journal on Information Security 2011, 1–25, 2011

[3]. S. B. ZANNOU, T. Djara, A. Vianou, "Secured revocable contactless fingerprint template based on center of mass", 2019 3rd International conference on Bio-engineering for Smart technologies (BioSMART), 2019.

[4]. J. Breebaart, B. Yang, I. Buhan-Dulman and C. Busch. Biometric template protection. Datenschutz und Datensicherheit-DuD 33, 299–304, 2009.

[5]. S. Chikkerur, N. Ratha, J. Connell, R. Bolle, Generating registration-free cancelable fingerprint templates, in: 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems, pp. 1-6, 2008.
[6]. W. Yang, J. Hu, S. Wang, J. Yang, Cancelable Fingerprint Templates with Delaunay Triangle-based Local Structures, Cyberspace Safety and Security, Lecture Notes in Computer Science. 8300 pp. 81-91, 2013.
[7]. S. Tulyakov, F. Farooq, P. Mansukhani, V. Govindaraju, Symmetric hash functions for secure fingerprint biometric systems, Pattern Recognition Letters. 28 pp. 2427-2436, 2007.
[8]. T. Ahmad, J. Hu, S. Wang, Pair-polar coordinate-based cancelable fingerprint templates, Pattern Recognition. 44 (10-11) pp. 2555-2564,2011.
[9]. P. Das, K. Karthik, B.C. Garai, A robust alignment-free fingerprint hashing algorithm based on minimum distance graphs, Pattern Recognition. 45 (9) pp. 3373-3388, 2012.
[10]. W.J. Wong, A.B.J. Teoh, M.L.D. Wong, Y.H. Kho, Enhanced multi-line code for minutiae-based fingerprint template protection, Pattern Recognition Letters. 34 pp. 1221-1229, 2013.
[11]. C. Li, J. Hu, Attacks via record multiplicity on cancelable biometrics templates, Concurrency and Computation: Practice and Experience. 26 (8) pp. 1593-1605, 2013.
[12]. Song Wang, Wencheng Yang and Jiankun Hu, Design of Alignment-Free Cancelable Fingerprint Templates with Zoned Minutia Pairs Pattern Recognition 2017, http://dx.doi.org/10.1017/j.patcog.2017.01.019
[13]. Djara, ASSOGBA, NAIT ALI and Vianou, 'Fingerprint Registration Using Zernike Moments : An Approach for a Supervised Contactless Biometric System', (9), pp. 254–271, 2009.
[14]. Dubuisson, M. et al. DISTANCE Between Point Sets Research supported by a ', pp. 566–568, (1994).
[15]. Syed Sadaf Ali, Surya Prakash, 3-Dimensional Secured Fingerprint Shell, Pattern Recognition Letters, 2018, doi:10.1018/j.patrec.2018.04.017
[16]. A. A. Taha and A. Hanbury. An Efficient Algorithm for Calculating the Exact Hausdorff distance. IEEE Trans. on PAMI 37, 2153–2163, 2015.
[17]. Moujahdi, C. et al. 'Fingerprint shell: Secure representation of fingerprint template', Pattern Recognition Letters, 45(1), pp. 189–196, 2014. doi: 10.1016/j.patrec.2014.04.001.
[18]. Ali S. S. and Prakash, S. 'Enhanced fingerprint shell', 2nd International Conference on Signal Processing and Integrated Networks, SPIN 2015, pp. 801–805, 2015. doi: 10.1109/SPIN.2015.7095438.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 135-141 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Design and Implementation of State-PID Feedback Controller for Poultry House System: Application for Winter Climate

Ilyas Lahlouh*, Ahmed Elakkary, Nacer Sefiani

Mohammed V University of Rabat, High School of Technology of Sale (EST), LASTIMI Laboratory, System Analysis, Information Processing and Industrial Management, 11030, Morocco

ARTICLE INFO

Article history:
Received: 08 September, 2019
Accepted: 04 January, 2020
Online: 22 January, 2020

Keywords:
State-PID feedback controller
Poultry house
Disturbance
Stability analysis

ABSTRACT

The poultry house is the area where the chickens are maintained for the main purpose to improve the productivity and the environmental conditions for the broilers. The growth of chickens inside the poultry house can be affected by several criterions such as relative humidity and temperature. In this paper, we propose to design and implement experimentally a state-PID feedback controller in order to achieve a high stabilization of the dynamics systems of the poultry process. The purpose of this study is to keep the temperature and relative humidity at desired values and to eliminate the disturbance generated during the winter climate. The effectiveness of the proposed controller is evaluated through numerical examples and with an experimental poultry house prototype. To feature more the efficiency of the proposed approach implemented, a comparative study has been conducted between the results achieved by the proposed method and the state-PI feedback controller and that of a state feedback controller.

1. Introduction

The poultry house climate is defined as the principal part that can straightforwardly influence the efficiency and comfort of the raising birds. Notwithstanding, the control of the housed livestock building remains a confounded errand because of the nonlinear behavior of psychrometric mechanisms required on both physiological and psychological components.

In most cases, the poultry farming is mainly affected by the outside condition (external relative humidity and temperature) and by the control of the actuators of the system (PAD cooling, heating, and ventilation system) [1]. Recently, the agricultural sector has extended altogether regarding control and optimization strategies for the primary parameters influencing good livestock management.

The hygro-thermal parameters are the most critical environmental variable on the poultry as announced in [2], [3]. Additionally, it has mutually been shown that the relative humidity and temperature affect intently the animal's improvement in Morocco because of the dominating weather and the country's location.

In this regard, we target our works to extend the ongoing studies by designing a suitable control strategy that is created depending on the mathematical model of broiler houses under cold conditions.

Over the most recent couple of years, the late research investigation demonstrates a tendency to the utilization of progressively different advanced procedures for control systems and simulation for a livestock building.

For the most part, it exists two types of control techniques, traditional techniques, and modern techniques. In such manner, numerous researches on livestock building concerning climate-controlling have been suggested in the literature, for example, nonlinear robust control for pig house [4], nonlinear adaptive control for animals house [5], predictive control for temperature and humidity in a normally ventilated structure [6], SMC (Sliding Mode Control) to control both of interior temperature and relative humidity in an evaporative cooling system [7].

Others studies related to the optimization control have been developed in the livestock building, for instance, the application of

*Corresponding author: Ilyas Lahlouh, High School of Technology of Sale, +212676640500, Email: ilyaslahlouh@research.emi.ac.ma

the ant colony optimization (ACO) for control broiler house under hot climates [8] and the utilization of the particle swarm optimization (PSO) in the managing of the thermal requirement of the poultry house.

In other hands, the fuzzy logic controller has been well tested in a greenhouse and poultry house [9], authors are reasoned that is fuzzy controller was able to give a superior efficiency in the stability. Other researchers [10] and [11] have linked a PID with fuzzy controllers to control the temperature indoors poultry house, they had concluded that the fuzzy-PID structure has excellent performance over a PID controller as far as settling time and steady-state error. A comparison between fuzzy and on/ off controllers have been analyzed in [12] to test and keep up the ambient parameters inside a model poultry house. Accordingly, all these advanced control systems have demonstrated their similarity to be used on control structure applications for poultry houses.

In this paper, the present work describes an extension of the original work reported in Proceedings of the International Conference on Automation and Computing [13].

The contributions of this article are portrayed as looks for: (1) to design a state-PID feedback controller to guarantee nominal stability and to regulate the hygro-thermal parameter of the system to a given set point value. (2) Implement the proposed controller in a real poultry house prototype system.

Hence, the inspiration for the state-PID feedback controller in this manuscript originates from some papers dealing with a wide variety of systems such as the SISO and MIMO systems [14] ,[15],[16].

The control configuration begins with the deduction of the linearized hygro-thermal model of the poultry process proposed in [17], at that point develops the state-PID feedback controller utilizing the eigenvalue assignment algorithm.

The remainder of the paper is formulated as follows: the primary segment exhibits quickly the hygro-thermal model of the poultry process. The second section describes the methodology of the state-PID feedback controller (SFCPID). Finally, some simulation results and remarks for stabilization and disturbance rejection are given to show the effectiveness of the proposed controller.

2. Mathematical model

2.1. Hygro-thermal model

Considering the mathematical modeling of the previous work detailed in [17] and the proposed studies on climate modeling inside livestock buildings [6] [18], the hygro-thermal model administered the poultry house is:

$$\dot{x_1} = \frac{u_1}{a} - \frac{r}{a}u_2 + R_{cb}(T_b - x_1) - (\frac{K_g}{a} + \frac{u_3}{V_b})(x_1 - T_{ext}) \quad (1)$$

$$\dot{x_2} = \frac{u_2}{b} + R_{ev}(0,26x_1^2 - 6,46x_1 + 81,6) - \frac{u_3}{V_e}(x_2 - H_{ext})$$

Where:

$$a = \rho_{air} V_b C_p \qquad b = \rho_{air} V_e \qquad R_{ev} = \frac{N_c.0,001}{3600.b}$$

$$R_{cb} = \frac{N_c.0,081.M_i^{0,67}}{\rho_{air} V_b C_p (\overline{R_{coat(i)}} + \frac{1}{5,4 + \frac{0,052 + 1,55.\vartheta_i^{0,48}.M_i^{0,16}}{0,13.M_i^{0,33}}})}$$

The state space form of the model (1) is represented in equation (2). The study is conducted with the cold operations where the controller input related to the PAD cooling system isn't working ($E_{ev} = U_2 = 0$).

$$\begin{bmatrix} \dot{\delta x_1} \\ \dot{\delta x_2} \end{bmatrix} = \begin{bmatrix} -(R_{cb} + \frac{K_g}{a} + \frac{\overline{u_{3e}}}{V_b}) & 0 \\ R_{ev}(0,52\overline{x_{1e}} - 6,46) & -\frac{\overline{u_{3e}}}{V_e} \end{bmatrix} \begin{bmatrix} \delta x_1 \\ \delta x_2 \end{bmatrix} + \begin{bmatrix} \frac{1}{a} & -\frac{\overline{x_{1e}}}{V_b} \\ 0 & -\frac{\overline{x_{2e}}}{V_e} \end{bmatrix} \begin{bmatrix} \delta u_1 \\ \delta u_3 \end{bmatrix} + \begin{bmatrix} \frac{K_g}{a} + \frac{\overline{u_{3e}}}{V_b} & 0 \\ 0 & \frac{\overline{u_{3e}}}{V_e} \end{bmatrix} \begin{bmatrix} \delta d_1 \\ \delta d_2 \end{bmatrix} \quad (2)$$

$$\begin{bmatrix} \delta Y_1 \\ \delta Y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \delta x_1 \\ \delta x_2 \end{bmatrix}$$

Where x are states variables associated with the internal relative humidity and temperature, u_1 and u_2 are the input controller associated with the heating system and the ventilation system, d are external disturbances.

The linearized model (2) can be simplified and represented to as a state space structure yielded (3):

$$\begin{aligned} \dot{\delta}_x &= A\delta_x + B\delta_u + B_d\delta_d \\ \delta\dot{y} &= C\delta x \end{aligned} \quad (3)$$

Where:
C is the output matrix, B and B_d are the input matrix and A is the state matrix.

2.2. Local Stability

We recommended concentrating to study the global stability of the general system (nonlinear) from its linearized tangent.

Let us consider the nonlinear system expressed as:

$$\begin{aligned} \dot{x} &= f(x) \\ f(x_e) &= 0 \end{aligned} \quad (4)$$

The linearized of the system around the operating point x_e is:

$$\delta\dot{x} = A(x_e)\delta x$$
$$\delta x = x - x_e$$

$$A(x_e) = \left(\frac{\partial f}{\partial x}\right)_{x_e} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & .... & \frac{\partial f_1}{\partial x_n} \\ . & ..... & . \\ . & & . \\ \frac{\partial f_n}{\partial x_1} & .... & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

After calculating, we get:

$$A(\overline{x_{1e}}) = \begin{bmatrix} -(R_{cb} + \frac{K_g}{a} + \frac{\overline{u_{3e}}}{V_b}) & 0 \\ R_{ev}(0{,}52\overline{x_{1e}} - 6{,}46) & -\frac{\overline{u_{3e}}}{V_b} \end{bmatrix}$$

If the case of $\overline{u_{3e}} = D_v = 0$, especially when the livestock building system work just with the heating system ($Q_T \neq 0$).

The eigenvalues of the matrix are:

$$\lambda_1 = 0 \qquad \text{and} \qquad \lambda_2 = -(R_{cb} + \frac{K_g}{a})$$

It is clear that the process is locally on the cutoff of the stability, and requires stabilization with an appropriate controller.

3. State-PID feedback controller design

In this section, state feedback and PID controllers are designed for the linear state space proposed model of the broiler house process.

3.1. State feedback controller

Let's consider the LTI system which having the following form:

$$\dot{x} = Ax + Bu \tag{5}$$

Where:

$A \in R^{nxn}$ $B \in R^{nxm}$ $x \in R^{nx1}$ $u \in R^{mx1}$

The linear system under consideration must to be totally controllable, the controllability matrix must have a full rank-n [19][20] and can be communicated as

$$C = [B \quad AB \quad A^2B \quad .... \quad A^{n-1}B] \tag{6}$$

The pair (A, B) is controllable.

Figure 1 illustrates the basic feedback controller of the poultry house.

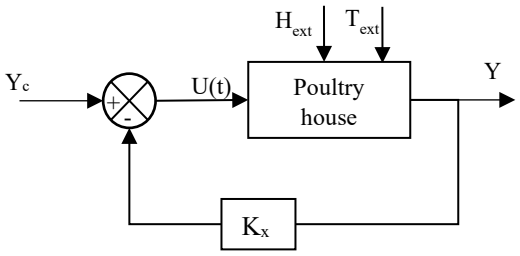


Figure 1: State feedback controlled poultry house system.

Where Y_c, U, Y and K_x are the command signal, the control law, the output of the plant and the feedback controller, respectively.

The control law can be composed on the structure:

$$u = Y_c - K_x x \tag{7}$$

The objective of this controller is to stabilize the system by the constant feedback control that imposes to place the poles at desired characteristic behavior for the states.

Using the feedback control law (7), the closed-loop characteristic polynomial system is given by:

$$|sI - (A - BK_x)| = 0 \tag{8}$$

The structure issue is to determine the feedback gain matrix K_x such that the closed-loop poles $\{\lambda_1, \lambda_2, ...\lambda_n\}$ of the system (8) are appointed at the desired values.

The methodology proposed to this problem is accomplished by using the pole placement algorithm.

The assignment placement algorithm concedes finding the matrix K_x, such as the closed-loop, system is stable. Both λ_i the desired closed loop poles and ν_i the eigenvector relating.

The following steps present the algorithm used for the assurance of the gain components of matrix K_x is: [21]

1. $(A - BK_xC)\nu_i = \lambda_i \nu_i$ Where : $K_x \in R^{mxn}$
2. $\nu_i = -(\lambda_i I - A)^{-1} BK_xC\nu_i$
3. $L_i = -(\nu_i I - A)^{-1} B$ With : $L_i \in R^{nxm}$

   $z_i = K_xC\nu_i$ With: $z_i \in R^m$
4. $z_i = (L_i^T R_i L_i)^{-1} L_i^T R_i \nu_i^d \quad i = 1,....,m$
5. $\nu_i^a = L_i z_i \qquad i = 1,....,m$

   $J_i = \frac{1}{2}(\nu_i^d - L_i z_i)^T R_i (\nu_i^d - L_i z_i)$
6. $K_x = ZV_a^{-1}$

With: $Z = [z_1 \quad z_2 \quad ... \quad z_m]$ and

$V_a = [\nu_1^a \quad \nu_2^a \quad ... \quad \nu_m^a]$

3.2. State-PID feedback controller

In this area, we introduce the proposed strategy control; we use in conjunction the state feedback and the PID controller.

The Proportional Integrator Derivative structure is one of the most controller broadly utilized in the industry, covering more than 90% of the industrial need in control. This predominance is because of its simplicity to implement it in a large plant does not require detailed knowledge. The control block scheme is shown in figure 2.

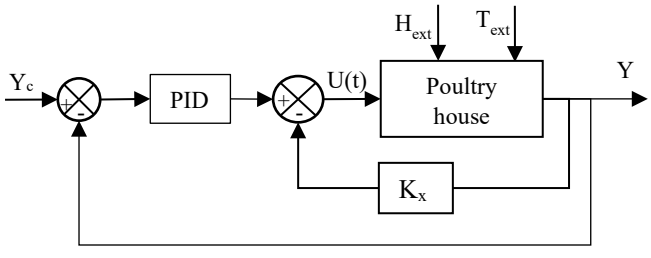


Figure 2: State-PID feedback proposed controller.

The control law u of the state-PID feedback proposed controller is :

$$U = K_p e + K_i \int e(t)dt + K_d \frac{de(t)}{dt} - K_x x \qquad (9)$$

Where $(K_p, K_i, K_d) \in R^{nxm}$ are the planned gain matrices to achieve a desired closed-loop system.

In summary, the problem described is resolved for a pole placement method consisting of three separated steps as mentioned in Guo et al [22].

The pole placement by state-P feedback is directed first prompting an intermediate system. Furthermore, state-I feedback is executed; another transitional system resulted. At last, the closed-loop system with the desired characteristic polynomial is acknowledged by state-D feedback joined as the final plan configuration.

The components of state-PID controller gain matrices for the PID structure can be obtained using the transformation into Frobenius canonical form detailed in [23],[24] and the equations mentioned in [15].

4. Results and discussions

In this analysis, we suggest regulating the internal temperature T_{int} and relative humidity H_{int} under cold climates, for that we design the state-PID feedback controller to get synchronization to deal with both ventilation and heating systems.

To affirm the performance of the proposed state-PID feedback controller, several comparative illustrative simulations are delineated.

4.1. Case 1: simulation study

The industrial poultry house system was equipped with controller natural ventilation, heating and PAD cooling system, the livestock building dimensions were 12,4m x 120m x 3,85m.

The real plant data have been saved for the entire period for the chickens raising of [20-21] days under the climates of the Mediterranean region (Rabat area, Morocco). (Figure 3)



Figure 3: Poultry house system used in the case study.

Table 1 presents the numerical estimations of the distinct criterions used in the simulation of the proposed controller.

Table 1: Input specifications used in the hygro-thermal model

Specifications	Values	Specifications	Values
ρ_{air}	1,2kg/m^3	T_b	41°C
C_p	1006J/kg.K	N_c	24000
V_b	3033m^3	V_e	3033m^3
M	0,905kg	ϑ	0,2m/s
$\overline{R_{coat}}$	0,2 m^{2}K/W	K_g	3126W/°C

Design of state feedback controller

The conventional (SFC) design technique begins with the determination of the desired eigenvalues from the given specifications of the controller K_x : λ_1^d , λ_2^d of an order of multiplicity respectively (n) and (n-1).

The basis of controllability is verified of the system (2) after including the different parameters:

Rank (C) =2

The specifications considered for the poultry house system are selected primarily on diminishing the steady-state error and the settling time.

After choosing the desired pole and applying the assignment eigenvalues algorithm, we find:

$$K_x = 10^3 \begin{bmatrix} 3659 & -51282 \\ -0 & -2 \end{bmatrix}$$

The input control of the system is :

$$u = 10^3 \begin{bmatrix} -3659 & 51282 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} T_{int} \\ H_{int} \end{bmatrix}$$

Design of state-PID feedback controller

The feedback gain matrix is computed as:

$$K_x = 10^3 \begin{bmatrix} 25637 & -76923 \\ -0 & -3 \end{bmatrix}$$

The PID gain matrices is obtained as :

$$K_p = 10^3 \begin{bmatrix} 3663 & -51282 \\ -0 & -2 \end{bmatrix} \qquad K_i = 10^3 \begin{bmatrix} 43956 & -51282 \\ -0 & -2 \end{bmatrix}$$

$$K_d = 10^3 \begin{bmatrix} 25637 & -76923 \\ -0 & -3 \end{bmatrix}$$

Stability analysis

The comparison of the closed loop response of the hygro-thermal parameters concerning the desired values of poultry house system with conventional state feedback, state-PI feedback and state-PID feedback controller using the package Simulink of MATLAB are demonstrated in figure 4 and 5.

For numerical simulation, the desired relative humidity and internal temperature have been fixed at 66% and 26°C.

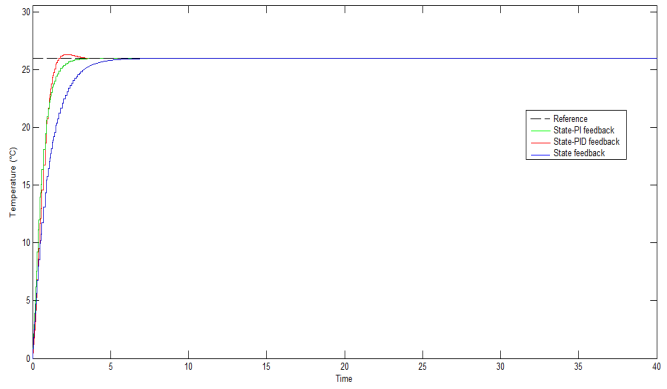


Figure 4: Temperature response generated by the different controller.

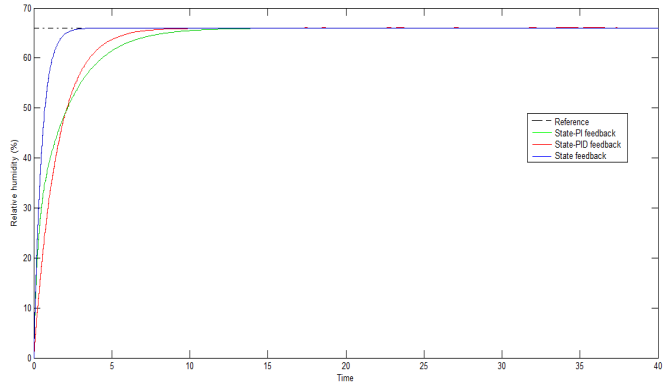


Figure 5: Humidity response supplied by the different controller

The state-PID feedback controller configuration performed wells; however, the state feedback controller performs moderately. It can be clearly seen in figure 4 and 5 that there are different responses in the transient state performance. The numerical values of these performances of the closed loops are tabulated in table 2 for comparison.

Table 2: Numerical values of standard performance indices.

	Performances indices	SFC	SFCPI	SFCPID
Temperature	Rise time (s)	2,42	1,38	1,21
	Settling time (s)	3,128	1,66	1,49
	Overshoot (%)	0	0	1,54
Relative humidity (%)	Rise time (s)	1,12	4,14	3,76
	Settling time (s)	1,31	5,62	4,37

From table 2, it appears that the state feedback controller gives poor results compared to that of state-PI and PID feedback controller in control of temperature. Other hands, the SFCPI gives a slow response compared to that of state-PID feedback state and a state feedback controller. Furthermore, the SFCPID affords the littlest value of rise time evaluated at 1,21s and the smallest value of settling time estimated at 1,49s.

It tends to be additionally observed that the performances indicated by the proposed controller are closely at their best values. The transient responses of states are plotted in figure 6 and 7.

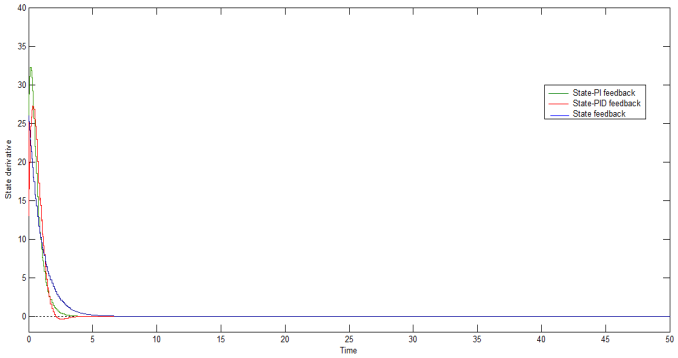


Figure 6: Response of the state derivative of the internal temperature.

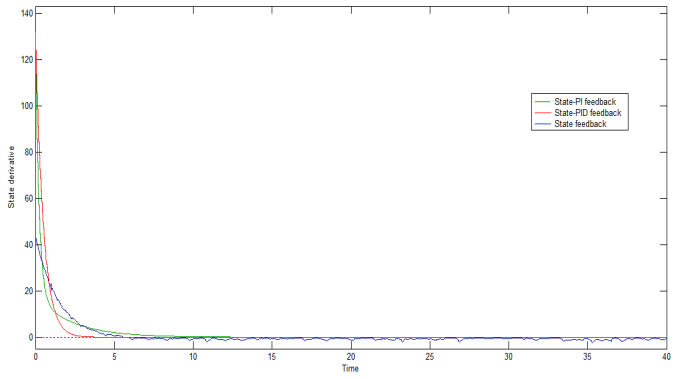


Figure 7: Response of the state derivative of the relative humidity.

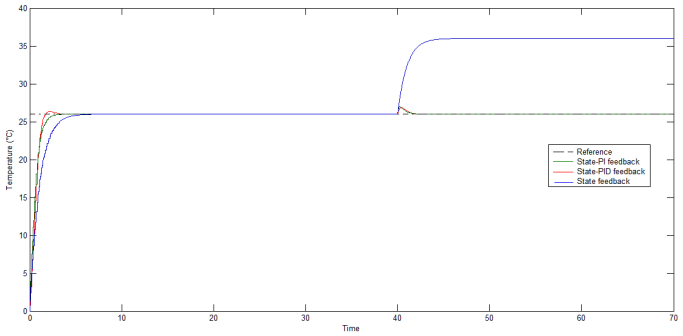


Figure 8: Temperature response simulated under disturbance with the different controller.

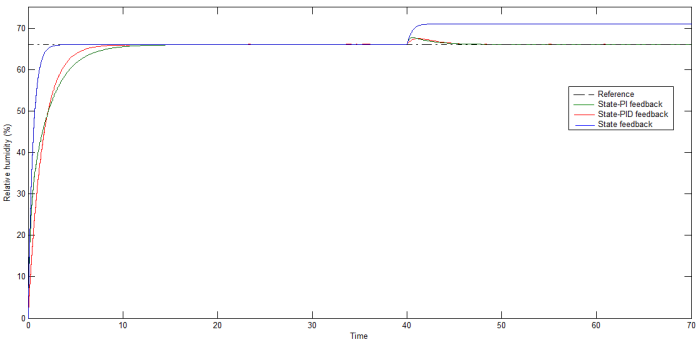


Figure 9: Relative humidity response simulated under disturbance with the different controller.

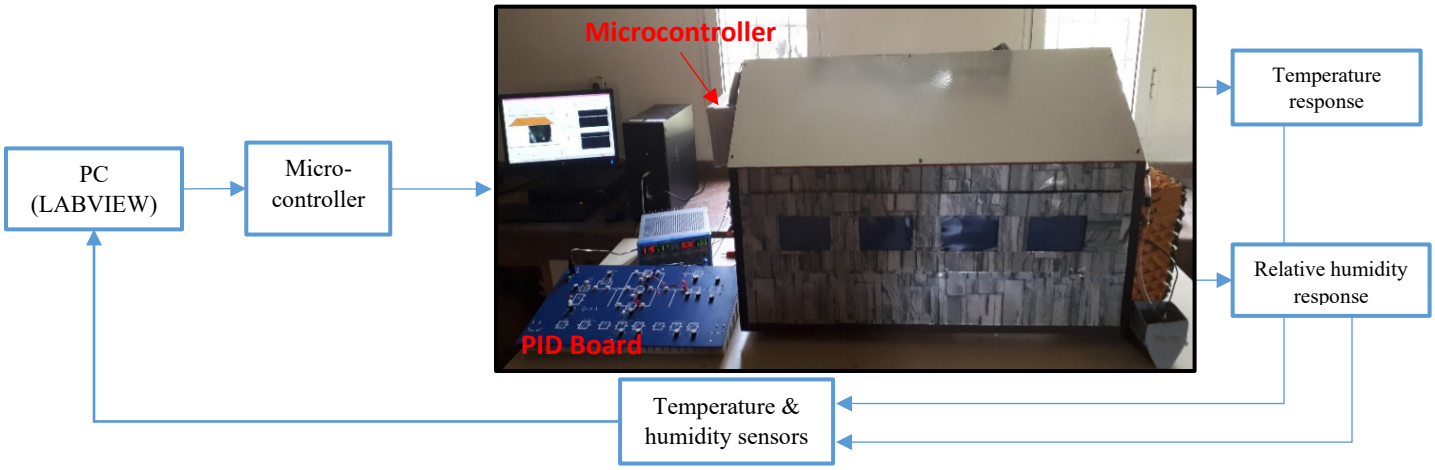


Figure 10: Experimental poultry house prototype.

From figures 6 and 7, it can be noticed that the derivative of the internal temperature accelerates rapidly within the state-PID feedback controller compared to that of SFCPI and that of state feedback. Also, it can be observed that the derivative response converges to zero after a short duration and this proves that the internal temperature and relative humidity have been well stabilized.

Disturbance rejection

The temperature and humidity response under the external disturbance with the three controllers are displayed in figures 8 and 9.

From figures 8 and 9, it very well may be obviously observed that with our proposed controller that the system reactions quicker, and recuperates totally from the external disturbance without any residue errors. This is because of a derivative and integral parameters acting as an error corrector.

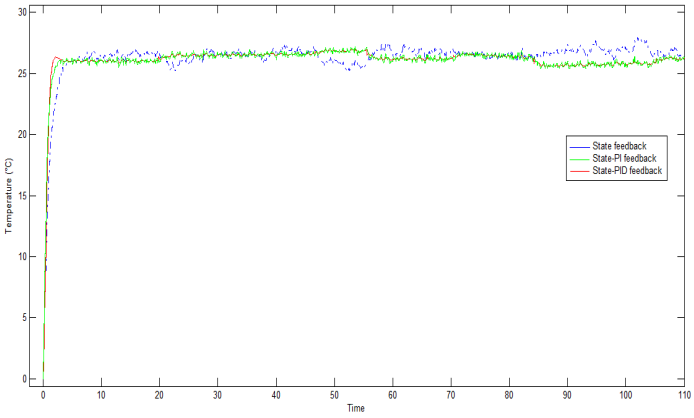
Furthermore, it tends to be seen that the planned state-PID feedback controller maintains the internal temperature and relative humidity within a short time of 7 seconds after applying a disturbance; however, the conventional state feedback controller isn't able to reduce or eliminate the external disturbance in both of temperature and relative humidity response.

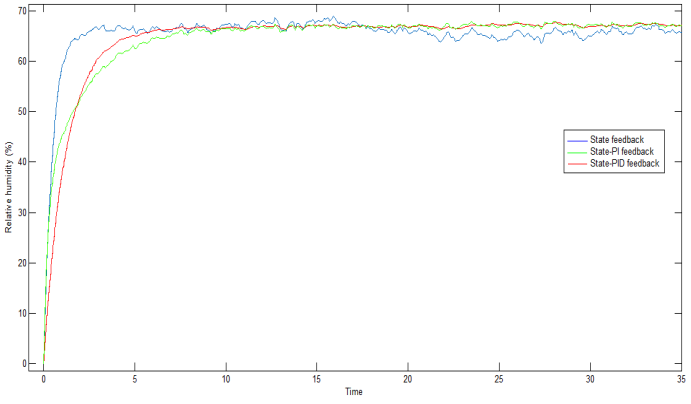
4.2. Case 2: Experimental poultry house prototype

The poultry house prototype has two main actuators, namely, the electric heater and the fans systems. Other subsystems contain lighting and evaporative cooling systems are also installed. The system program includes the main algorithm unit for all systems in the broiler house. In addition, the experimental data is transmitted to the computer device using a micro controller. The real data is monitored with the LabVIEW software; the users can control the important operations and can to be informed of the various changes for the system. (See figure 10)

The internal temperature is measured using LM35 and the relative humidity is acquired via HIH4030 sensor. The variable input of heating and ventilation system are fitted with a thyristor driver circuit in order to produce a varying voltage for running the fans and the heater at different level speed and stage. The change in applied voltages is legitimately relative to the airflow speed and heating rate.

The three controllers are implemented for the poultry house prototype and experimental results of the temperature and relative humidity response are represented in figures 11 and 12.





The implementation of the different controller to the poultry house prototype control system show the better quality of using SFCPID in the regulation of the system. The results obtained with

the state-PID feedback controller provide a better response for relative humidity and temperature response with a very smooth variation in tracking the desired set points.

5. Conclusions

In this article, we have proposed and designed a state-PID feedback controller for stabilizing the relative humidity and temperature indoor the broiler house process during the colds conditions.

Numerical examples are tested to confirm the effectiveness of the proposed design controller to stabilize the system. On the other, we have considered the existence of disturbance simulated as the external perturbations getting from the outside climate. The results of simulations show that the proposed methodology was proficient regarding dynamic performance, for example, a decrease of settling time and rise time. Moreover, the SFCPID was capable of tracking the desired set point with a rapid rate of convergence. Thus, the proposed controller has proven its performance through the simulations under disturbance and the comparative study conducted between the state feedback and state-PI feedback technique and the state-PID feedback controller.

In addition, the proposed controller is implemented for the real experimental poultry house prototype. It is concluded that the proposed state-PID feedback controller affords excellent performance than the state-PI feedback and state feedback controller in terms of following the set points parameters.

Finally, it is pointed out that the proposed control methodology of the poultry house accentuates the regulation of the system and the optimal parameters.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by the National Center for Scientific and Technical Research (CNRST) within the context of the Research Excellence Awards program.

References

[1] S. Yahav, A. Straschnow, D. Luger, D. Shinder, J. Tanny, and S. Cohen, "Ventilation, sensible heat loss, broiler energy, and water balance under harsh environmental conditions," *Poult. Sci.*, vol. 83, no. 2, pp. 253–258, 2004.

[2] N. J. Daghir, Ed., *Poultry production in hot climates*, 2nd ed. Wallingford: CABI, 2008.

[3] T. Banhazi *et al.*, "Issues related to livestock housing under hot climatic conditions including the animals' response to high temperatures," *Anim. Hous. Hot Clim.*, p. 4, 2007.

[4] A. G. Soldatos, K. G. Arvanitis, P. I. Daskalov, G. D. Pasgianos, and N. A. Sigrimis, "Nonlinear robust temperature–humidity control in livestock buildings," *Comput. Electron. Agric.*, vol. 49, no. 3, pp. 357–376, Dec. 2005.

[5] P. I. Daskalov, K. G. Arvanitis, G. D. Pasgianos, and N. A. Sigrimis, "Non-linear Adaptive Temperature and Humidity Control in Animal Buildings," *Biosyst. Eng.*, vol. 93, no. 1, pp. 1–24, Jan. 2006.

[6] P. I. Daskalov, "Prediction of temperature and humidity in a naturally ventilated pig building," *J. Agric. Eng. Res.*, vol. 68, no. 4, pp. 329–339, 1997.

[7] K. Senawong, S. Winitchai, and T. Radpukdee, "Humidity and temperature control in an evaporative cooling system of a poultry house," *Eng. Appl. Sci. Res.*, vol. 39, no. 1, pp. 95–111, 2012.

[8] I. Lahlouh, A. El Akkary, and N. Sefiani, "PID/Multi-Loop Control Strategy for Poultry House System Using Multi-Objective Ant Colony Optimization," *Int. Rev. Autom. Control IREACO*, vol. 11, no. 5, p. 273, Sep. 2018.

[9] K. Chao, R. S. Gates, and N. Sigrimis, "Fuzzy logic controller design for staged heating and ventilating systems," *Trans. ASAE*, vol. 43, no. 6, p. 1885, 2000.

[10] D. O. Aborisade and O. Stephen, "Poultry house temperature control using Fuzzy-PID controller," *IJETT Chennai*, vol. 11, no. 6, pp. 310–314, 2014.

[11] R.Manoj Manjunath and S.Janaki Raman, "Fuzzy adaptive PID for Flow Control System based on OPC." IJCA Special Issue on "Computational Science-New Dimensions & perspectives," 2011.

[12] E. Mirzaee-Ghaleh, M. Omid, A. Keyhani, and M. J. Dalvand, "Comparison of fuzzy and on/off controllers for winter season indoor climate management in a model poultry house," *Comput. Electron. Agric.*, vol. 110, pp. 187–195, Jan. 2015.

[13] I. Lahlouh, A. El Akkary, and N. Sefiani, "State feedback controller in a closed poultry house system," in *24th International Conference on Automation and Computing*, Newcastle upon Tyne, 2018.

[14] W. Wiboonjaroen and S. Sujitjorn, "State-PID Feedback for Magnetic Levitation System," *Adv. Mater. Res.*, vol. 622–623, pp. 1467–1473, Dec. 2012.

[15] S. Sujitjorn and W. Wiboonjaroen, "State-PID Feedback for Pole Placement of LTI Systems," *Math. Probl. Eng.*, vol. 2011, pp. 1–20, 2011.

[16] I. H. Akyuz, E. Yolacan, H. M. Ertunc, and Z. Bingul, "PID and state feedback control of a single-link flexible joint robot manipulator," in *2011 IEEE International Conference on Mechatronics*, 2011, pp. 409–414.

[17] I. Lahlouh, A. El Akkary, and N. Sefiani, "Mathematical Modelling of the Hygro-Thermal Regime of a Poultry Livestock Building: Simulation for Spring Climate," *Int. Rev. Civ. Eng. IRECE*, vol. 9, no. 2, p. 79, Mar. 2018.

[18] International Commission of Agricultural Engineering, Ed., *CIGR handbook of agricultural engineering*. St. Joseph, MI: American Society of Agricultural Engineers, 1999.

[19] G. Conte, C. . Moog, and A. M. Perdon, *Algebraic methods for nonlinear control systems*. London: Springer-Verlag London Limited, 2007.

[20] P. J. Antsaklis and A. N. Michel, *Linear Systems*. Boston (MA): Birkhäuser, 2006.

[21] R. Fonod and P. Kocsis, "State Feedback Control Design Using Eigenstructure Decoupling," in *18th International Conference on Process Control*, 2011, pp. 268–272.

[22] G. Ge, M. Zi, and Q. Junfei, "State-PID feedback control with application to a robot vibration absorber," *International Journal Modelling, Identification and Control*, 2006.

[23] D. Luenberger, "Canonical forms for linear multivariable systems," *IEEE Trans. Autom. Control*, vol. 12, no. 3, pp. 290–293, Jun. 1967.

[24] M. Valášek and N. Olgac, "Efficient eigenvalue assignments for general linear MIMO systems," *Automatica*, vol. 31, no. 11, pp. 1605–1617, 1995.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 138-142 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

The Acquiring Process of Entrepreneurship Competencies in Realist Airbrush Art

Bayu Rahmat Setiadi[1,*], Suparmin[2], Slamet Priyanto[2], Henny Pratiwi[1], Sugiyono[1]

[1]*Department of Mechanical Engineering Education, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia*

[2]*Department of Mechanical Engineering Education, Universitas Sarjanawiyata Tamansiswa, Yogyakarta, Indonesia*

ARTICLE INFO

Article history:
Received: 14 October, 2019
Accepted: 05 February, 2020
Online: 09 March, 2020

Keywords:
Realist Airbrush
Art and Technologies
Acquiring Process
Vocational High School

ABSTRACT

One of the creative industries with rapid growth is the airbrush industry. Airbrush art, especially realist, is identical to the field of automotive engineering in the combination of art and technology. The research is directed to prove how the difficult realist airbrush competency can be achieved by Vocational High School (VHS) students in a short time. Also, the research synchronises the increase in entrepreneurial understanding of VHS students in real business incubation in the realist airbrush industry. The final results show that entrepreneurial learning needs to add elements of art as a stage in the acquisition of real competencies.

1. Introduction

The creative industry is one of the creative economies that can stimulate the economy in Indonesia. Some of these creative sectors have great opportunities in solving the problem of unemployment in Indonesia [1]. Creative industries support almost half of Indonesia's foreign exchange [2]. Data shows that the number of entrepreneurs from the creative industries is 49% spread widely in some province in Indonesia where Yogyakarta Province as the centre [3].

Creative industries that grow in Indonesia have several sub-sectors include animation, architecture, design, photography, music, crafts, culinary, fashion, research and development, publishing, film, advertising, interactive games, performance art, visual arts, information technology, television, radio, and video. Some of these sectors, six sub-sectors with the highest growth, are television, film-animation-video, performing arts, crafts, and visual communication design [4]. One of the most desirable sub-sectors is craft and art [5]. Yogyakarta City is one of the contributors to the arts and crafts business in regional, national, and international export marketing [6].

The growth of the creative arts industry in Yogyakarta is a reflection of artists in Indonesia as inspiration for ideas and creativity. One of the trendy creative endeavours in this city is airbrush art. The famous airbrush in this city is Planet Airbrush Potlot. Potlot has been established since 2000 and has produced hundreds of airbrush works on automotive, painting, wall, and other media. Potlot won various international events. According to an art perspective, the realist airbrush technique is very complicated and is in great demand by many people because of the high level of detail and animate the character of the artwork [7]. So, Yogyakarta could be said as a place for artists in exploring art [8].

The effort to open a realist airbrush, as exemplified in the previous paragraph, explains that there are good prospects for Vocational High Schools (VHS) to produce graduates as airbrush entrepreneurs [9]. The trend of shifting from VHS graduates to entrepreneurs is due to the limited number of jobs for VHS graduates to get a new job. However, leading graduates to create airbrush art is an option to reduce the unemployment level of VHS graduates.

Based on the previous description, the researcher wants to examine how VHS students who are doing internships in the realist airbrush industry can be skilled in simple airbrushing. Besides, observing the process of acquiring entrepreneurial competencies about the business is the key to this paper because the research is unique and can be an inspiration for business people to be interested in opening up new business fields.

2. Combining Art in Automotive for New Business Incubation

2.1. Collaboration between creative industry subsector in Indonesia

* Bayu Rahmat Setiadi, Email: bayursetiadi@uny.ac.id

The creative industry in Indonesia has a variety of sub-sectors. The Creative Economy Agency or abbreviation of Bekraf Indonesia has released 16 strategic sub-sectors, including animation, architecture, design, photography, music, crafts, culinary, etc. Some of these sub-sectors were developed in several iconic big cities such as Yogyakarta (handicraft city), Pekalongan (city of Batik), Bandung (fashion city) [10]. The iconic cities became a centre for business development and creative industries as well as references from various regions [11].

The collaboration between various industry sub-sectors is a good strategy in order to produce quality products. The products currently available are a combination of multidisciplinary or better known as Science, Technology, Engineering, Arts, and Mathematics (STEAM) [12]. Science and technology need to work together with art to manufacture products that are more accepted by society. It indicates that the creative industries sub-sectors need to collaborate dynamically and flexibly so that the products in the creative industries become more diverse and quality.

2.2. Prospective realist airbrush as a promising business

Fine art has two types, namely, pure and applied art [13]. Pure art, or better known as first works, is an expression of the soul that is etched through dots, lines, paintings, and other forms of images. Applied art is a form of application of artists to the media that it has designed. Airbrush, as an art-based painting technique, is an applied art that is expressively determined by painting in the media [14].

Realist airbrushes have advantages in producing real artwork. The examples are such as in murals, posters, contemporary, and painting [15]. Not many people are experts in painting realists in certain media. A high level of difficulty in drawing and depth in inspiration determines the success of a realist. Also, the airbrush technique requires much experience so that the results are more realistic and easily understood by connoisseurs of art [16]. Various challenges in airbrush art make the techniques needed in the business of painting and airbrush. Therefore, a realistic airbrush with complicated media such as vehicles, motorcycles, cars, trucks, and buses is an excellent business opportunity.

2.3. VHS automotive students have potential in the realist airbrush industry entrepreneurship

The curriculum determines VHS graduates to work, go to college, and be entrepreneurs [17]. The current curriculum design in Indonesia refers to the 2013 Curriculum. The contents of the curriculum with the priority of practicum indicate a change in mindset for VHS to increase psychomotor ability compared to cognitive and adaptive. For example, automotive skills in the curriculum spectrum in Indonesia limit the practice of basic painting. If this continues, the tendency of students to get late for the first job becomes a common problem in Indonesia. The entrepreneurship is like a medicine needed in solving the unemployment problem in Indonesia.

Policies in Indonesia encourage entrepreneurship in vocational education. Many studies have explored the opportunities for creative industry employment as a potential face of new entrepreneurs. Cheap, collaborative capital, a wide variety of products, and the dynamics of consumer change are ideal potentials in entrepreneurial learning. In the case of automotive engineering, the business of marketing and repair services are frequent, and consumers have regular customers who certainly reduce the interest of entrepreneurship for beginners.

3. Research Method

The study used a mixed-method with an explanatory cycle model. This method is in line with the process of acquiring competencies that have three observed research stages. The first step using phenomenology to reveal the process of achieving student competence in airbrush techniques. The next step is adopting experimentation in the form of pre-test and post-test to examine the first hypothesis. The last step is testing the second hypothesis with pre-test and post-test in order to validate the improvement of entrepreneurship skills in airbrush realism after the incubation process was carried out.

The population is 30 students at VHS Ma'arif-Salam, Magelang Regency, majoring in automotive engineering. Data collection techniques used were interviews and observations and questionnaires. Interviews were conducted to obtain in-depth information from the instructor. Meanwhile, the pre-test and post-test instruments will ask questions related to increasing understanding of both airbrush and entrepreneurship training — data analysis by reducing the critical information based on coding and besides, interpreting the findings through a crosscheck of interviews and observations to the instructor. Descriptive statistics were used to measure the understanding of both airbrush and entrepreneurship training.



Figure 1: The Art of Collaborating with Automotive

4. Results

The acquisition of entrepreneurial competence in the realist airbrush has two main scopes. They are achieving airbrush success and success in opening a business. The results of the two studies are mutually simultaneous to prepare VHS graduates to open the automotive business in the realist airbrush field. Based on observations, interviews, and data reduction, the flow chart as obtained for VHS students who were said to be amateur workers.

Figure 2 describes the process of obtaining entrepreneurial competence in airbrush skills. The process illustrates that the repetition of the process occurs as a stage of student learning. For example, students are often less thorough in sanding and caulking. The results of sandpaper and picking should be checked for smoothness because it has an impact on the final product. Therefore, the initial work raises many related cycles as capital to prepare an applicable and neat airbrush media. The improvement of airbrush realist skills is carried out continuously by conducting tests at the beginning and providing posttest after treatment.

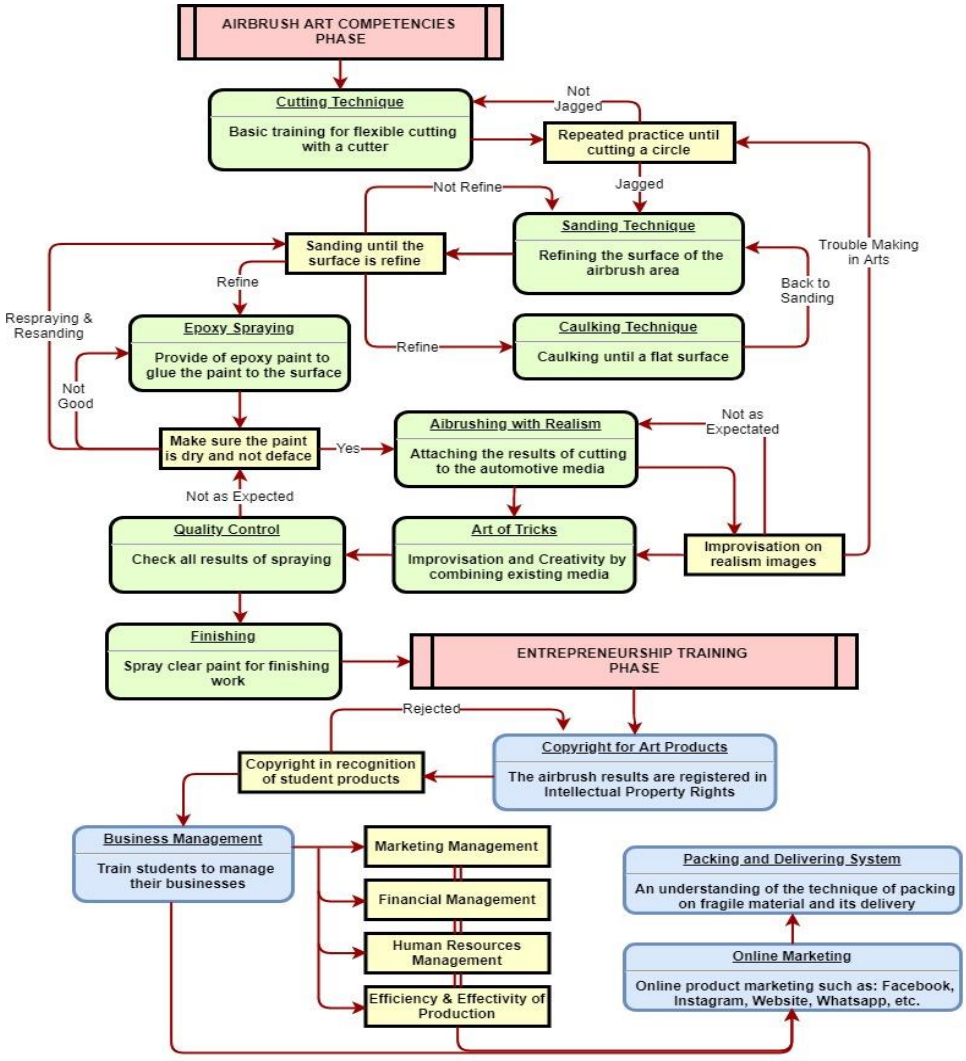
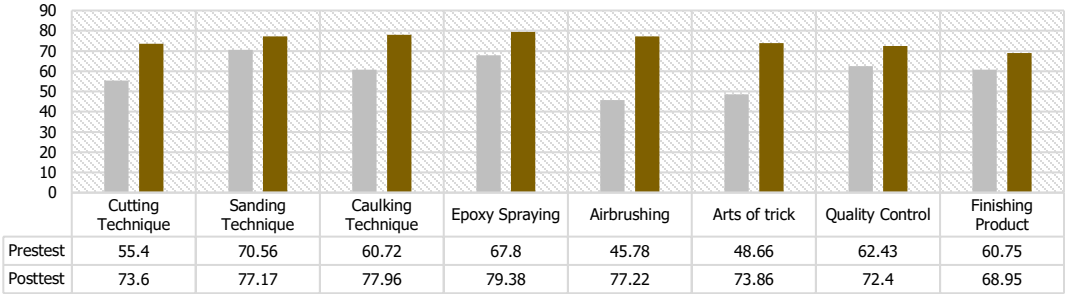


Figure 2: The Process of Forming Entrepreneurship Competence in the Automotive Realist Airbrush Business



	Cutting Technique	Sanding Technique	Caulking Technique	Epoxy Spraying	Airbrushing	Arts of trick	Quality Control	Finishing Product
Prestest	55.4	70.56	60.72	67.8	45.78	48.66	62.43	60.75
Posttest	73.6	77.17	77.96	79.38	77.22	73.86	72.4	68.95

Figure 3: Realist Airbrush Skills in Pretest and Posttest

There is a significant influence on students' skills due to airbrush training (Fig. 3). More than one-fifth of the posttest outperformed the pretest score, indicating that students' interest in painting was very high. Substantially, the enthusiasm of students to continue on the entrepreneurial step becomes an opportunity for them to be confident in opening a realist airbrush business. Positive responses to the creative industries in this sector need to be encouraged comprehensively.

The challenge in airbrush training based on interviews are cutting techniques, airbrushing, and arts of tricks. All three techniques require a lot of practice and material. These obstacles can be overcome if they often practice continuously to form these skills. Behaviorally, a child's skills will work well if he does repetitive work according to the object being worked on the workplace. Meanwhile, constructivism point of view states that learning can build independence in the field of arts, creation, and innovation.

The entrepreneurial skills of VHS students are measured by doing pretest and posttest before the airbrush training is carried out, and after they get the entrepreneurial knowledge. In the posttest, the skill of students increased by 11.38% compared to pretest. The following is a graph of the increasing entrepreneurial development of VHS students. The improvement of realist airbrush skills examination is carried out continuously by conducting tests at the beginning through a pretest and giving a posttest after treatment.

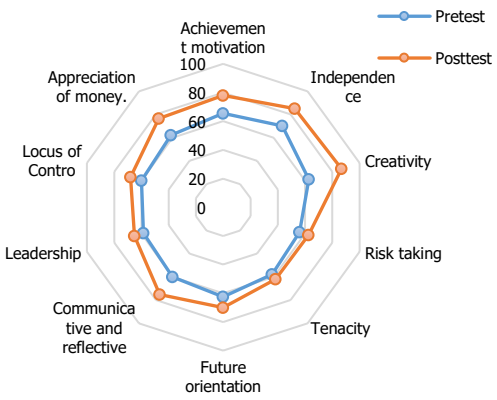


Figure 4. Distribution of VHS Student Entrepreneurship Experience Progress

The distribution of VHS student entrepreneurship experience progress diagram (Fig. 4) shows that there is an increase in all aspects of entrepreneurship criteria. The remarkable improvement is an increase in creativity, which is a quarter per cent of the value of the two tests. In the other side, there is a concern in the low level of tenacity. The short-term training and entrepreneurial experience have not significantly increased their tenacity level, but the entrepreneurship training after completing the product provides a significant increase in other criteria.

5. Discussions

Arts and automotive would be two beautiful fields of science if they integrated. Maeda proves that the existence of art in automotive design will increase the sale value of the vehicle [12]. A significant increase is very noticeable when automotive products begin to be marketed at certain specific events. The research findings show the acceleration of students' ability in the mastery of art in the automotive field on airbrush realism. The primary abilities of students become the first reason in increasing their motivation and interest to explore airbrush science based on the needs of automotive work.

Entrepreneurship that is in line with skills is an excellent opportunity to pursue the field of expertise [18]. Students with airbrush skills have a big challenge to open the airbrush business later after they graduate. This entrepreneurial theory will undoubtedly oppose the assumption of Prosser & Quigley theory that working in the automotive field is required to work in the automotive industry [19]. This assumption began to shift since Dewey gave the freedom to determine the goals and lives of students following the needs and opportunities of students [20]. However, this finding is part of providing information to schools in Indonesia that VHS has great potential to direct graduates to become entrepreneurs.

6. Conclusion

There is a significant influence on students' skills due to airbrush training. More than 20% of the posttest outperformed the pretest score, indicating that students' interest in painting was very high. The distribution of VHS student entrepreneurship experience progress diagram also shows that there is an increase in all aspects of entrepreneurship criteria. The remarkable improvement is an increase in creativity which increase 25% after the training.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Our thanks to the Directorate of Research and Community Service - Ministry of Research, Technology, and Higher Education of the Republic of Indonesia for funding research into the development of creative industries for VHS students. Researchers are grateful to all those who helped the research process.

References

[1] B. R. Setiadi and Suparmin, "Providing Space for Creative Industry Sub-Sectors to Meet the Labor Market for Vocational School Graduates," in Advances in Social Science, Education and Humanities Research, 2017, vol. 102, no. Ictvt, pp. 309–313.

[2] M. E. Pangestu, Ekonomi Kreatif Indonesia 2025, no. 5. Jakarta: Departemen Perdagangan RI, 2015.

[3] Bekraf, Opus creative economy outlook 2019. Jakarta: Badan Ekonomi Kreatif, 2019.

[4] S. R. Maryunani and I. R. Mirzanti, "The Development of Entrepreneurship in Creative Industries with Reference to Bandung as a Creative City," Procedia - Soc. Behav. Sci., vol. 169, no. August 2014, pp. 387–394, 2015.

[5] ICCI, "KADIN Dorong Ekonomi Kreatif Yogyakarta," Indonesian Chamber of Commerce and Industry, 2017. .

[6] DIY Tourism Office, "Statistik Kepariwisataan 2015," Yogyakarta,Indonesia, 2015.

[7] X. Wang, "Realist painting and its relationship to my creative practice Realist painting and its relationship to my creative practice," Edith Cowan University, 2010.

[8] M. D. Marianto, "Surrealist Painting in Yogyakarta," University of Wollongong, 1995.

[9] B. R. Setiadi, Suparmin, S. Priyanto, and P. Roniwijaya, "Tracking the potential of vocational high school students leading to eighteen strategic

creative industry sub-sectors in Indonesia," IOP Conf. Ser. Mater. Sci. Eng., vol. 535, no. 1, 2019.
[10] F. Z. Fahmi, S. Koster, and J. van Dijk, "The location of creative industries in a developing country: The case of Indonesia," Cities, vol. 59, pp. 66–79, 2016.
[11] M. B. Alexandri, E. Maulina, A. Chan, and R. W. S. Sumadinata, "Creative Industries : Existence of Arts Traditional Industries in Indonesia," Acad. Strateg. Manag. J., vol. 18, no. 1, pp. 1–12, 2019.
[12] J. Maeda, "STEM + Art = STEAM," STEAM, vol. 1, no. 1, pp. 1–3, 2013.
[13] C. P. Magill, "Pure and Applied Art: A Note on Johann Peter Hebel," 1957, pp. 183–188.
[14] J. Lu, C. Barnes, S. DiVerdi, and A. Finkelstein, "RealBrush: Painting with examples of physical media," ACM Trans. Graph., vol. 32, no. 4, 2013.
[15] A. Oliver, "Critical realism in contemporary art," University of Pittsburgh, 2014.
[16] M. A. A. Soltan, "An Investigation into the History of the Airbrush and the Impact of the Conservation Treatment of Airbrushed Canvas Paintings Mohamed," University of Northumbria, 2015.
[17] H. Retnawati, S. Hadi, and A. C. Nugraha, "Vocational high school teachers' difficulties in implementing the assessment in curriculum 2013 in Yogyakarta Province of Indonesia," Int. J. Instr., 2016.
[18] B. R. Setiadi, Suparmin, and Samidjo, "Preparing engineering students for entrepreneurial creative industries," Glob. J. Eng. Educ., vol. 20, no. 2, pp. 127–131, 2018.
[19] P. Sudira, "Indonesia vocational education praxis betwen," in Empowering Vocational Education and Training to Elevate National Economic Growth, 2014, pp. 190–200.
[20] J. Dewey, Democracy and Education. Hazleton: The Pennsylvania State University The, 2001.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 140-146 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A New Wire Optimization Approach for Power Reduction in Advanced Technology Nodes

Jalal Benallal[1], Lekbir Cherif[*,1,2], Mohamed Chentouf[2], Mohammed Darmi[1], Rachid Elgouri[3], Nabil Hmina[1]

[1]*Laboratory of Systems Engineering, National School of Applied Sciences, Ibn Tofail University, B.P 241, Kenitra, Morocco*

[2]*ICDS Department, Mentor Graphics, Rabat, 11100, Morocco*

[3]*Laboratory of Electrical Engineering & Telecommunication Systems, National School of Applied Sciences, Ibn Tofail University, Kenitra, B.P 241, Morocco*

ARTICLE INFO

Article history:
Received: 17 September, 2019
Accepted: 10 November, 2019
Online: 25 November, 2019

Keywords:
Integrated circuit conception
Physical design (EDA)
Physical implementation
Interconnect
Low power
Power optimization
Dynamic power
Total power

ABSTRACT

In advanced technologies nodes, starting from 28 nm to 7 nm and below, the power consumed of integrated circuits (ICs) becomes a big concern. Consequently, actual electronic design automation (EDA) tools are facing many challenges to have low power, reduced area and keep having required performance. To reach required success criteria, and because each picosecond and each picowatt counts, continuous development of new optimization technics is necessary. In this paper, we put to the experiment and analysis a new technic to reduce IC consumed power by optimizing its interconnections (nets). We propose an optimal algorithm and enhance it for a better compromise between having less consumed power and keep having a good design rout-ability. The new wire optimization technic based on an optimal choice of target nets for optimization: which is the list of nets consuming more than 80% of the total power in the interconnection without exceeding 30% of the total number of nets. Experiments on 14 test cases show an average total power saving of 5% on both dynamic and total power.

1. Introduction

Modern system on chip (SoC) and network on chip (NoC) circuits are known by the integration of complex interconnect IP which brings more difficulties for the timing closure especially with the 16 nm technology node (and below) [1]. In parallel with the circuit performance, new technology nodes allowed a very high transistors' integration to have more functionalities inside smaller die area [2], which brings many manufacturing challenges before having production circuits. On the other hand, modern designs are power-constrained (e.g., IoT, Automotive, Mobile) [3, 4]. Thus, electronic design automation (EDA) tools should have all necessary functionalities for a good compromise between required circuit performances, manufacturing, and power consumption.

In advanced technology nodes, wire capacitance has become a key challenge to design closure, and this problem only worsens with each successive manufacturing process mainly due to the minimum spacing between adjacent wires [5, 6]. Today, a physical implementation flow for the digital circuit should be able to play with multiple scenarios during routing to find the best compromise between timing, power, and rout-ability. [7] Details the importance of interconnect optimization and how its optimization is playing a pillar role in chip performance. Also, [8, 9] Present more routing closure challenges.

Related to the power optimization, at the physical implementation, many technics are used targeting both leakage and dynamic power. [10, 11] Gives some technics to reduce power on the cells' element. [12, 13] present new technics for power optimization on the design network. [14, 15] focus more on the technology ways to decrease total consumed power in the design.

This work introduces a new technique for power optimization during the physical implementation of an IC by optimizing the wire capacitance of its power critical interconnections (nets). We then propose a new solution that directly improves upon the original solution [16] construction by proposing an enhancement of the procedure that gets target nets for optimization. The new solution helps for having good route congestion overflow while

[*]Corresponding Author: Lekbir Cherif, Technopolis, Building B0, Sala El Jadida, Morocco, +212611396180, Email Lekbir.cherif@gmail.com

keeping the same power reduction gain. The solution approach achieves improvement for both objectives, having a maximum power gain by simple nets re-routing, and reduce the overflow for better design rout-ability.

Experimental results on 14 test-cases made with most advanced technologies nodes (28, 20, 16&7 nm) demonstrate that this technique achieves an additional average of 5% on total power by targeting nets consuming more than 80% of the power.

This paper makes the following contributions:

- Starting with the final result from paper [16], in this paper, we propose an enhanced procedure that gets fewer data nets as a target for optimization.
- The enhanced solution helps on having acceptable routing congestion for better routing capability and, at the same time, keep having the same good power reduction gain of 5% in the average.
- Proves the benefit of this transform on power reduction on data nets by experimenting with the new flow on 14 real designs made with the most advanced technology nodes 28, 20, 16 and 7 nm.

The remainder of this paper is organized as follows. Section 2 presents the power trend on a modern integrated circuit and briefly, review the results achieved in paper [16]. Section 3 presents the enhanced solution that gets fewer targets for better routing congestion overflow as a wire promotion power-aware method. Section 4 reports our experimental results, and Section 5 concludes the paper.

2. Power in digital ICs at physical implementation stage

2.1. Power calculation and the trend in advanced technology nodes

Equations (1), (2) and (3) summaries how the IC power can be calculated: [17-19]

$$P_{dynamic} = P_{internal} + P_{switching} \quad (1)$$

$$P_{internal} = \frac{1}{2}.toggle_rate.(E_{rise} + E_{fall}) \quad (2)$$

$$P_{switching} = \frac{1}{2}.toggle_rate.(C_{load}).V^2 \quad (3)$$

Where:

$\boldsymbol{E_{rise}}$ and $\boldsymbol{E_{fall}}$ are rise and fall energy;

toggle_rate is number of toggles per time-unit;

$\boldsymbol{C_{load}}$ is total wire capacitance;

V is the power supply voltage

During IC physical implementation, only a few parameters could be optimized; for example: reduce the wire capacitance $\boldsymbol{C_{load}}$ by making the interconnection wire-length shorter or by spreading. Swap cells with high internal or leakage power with lower cells' power.

In advanced nodes 28 nm and below, static power consumption represents less than 10%, on average, of the total power consumed by an IC. The dynamic circuit power is composed of the internal power that represents 20% to 30% and the switching power representing 70% to 80% of the overall consumption of a circuit [20, 21]. These numbers are proved in test-cases we are using for this paper: Table 1, shows a summary of their main characteristics. Figure 1, shows their average static and dynamic power repartition. Figure 2 shows the different average power components repartition. It is evident to see that by targeting all the "Data nets", we are targeting more than 50% of the total power for optimization.

Table 1: Design Specifications

Designs	Number of instances	Number of nets	Physical area (µm²)	Technology (nm)
Design#1	535815	566853	407658	20
Design#2	557185	588503	470164	20
Design#3	905670	1268272	1072040	7
Design#4	137950	221084	46975.9	7
Design#5	477969	546368	236706	16
Design#6	3695650	8402104	5949560	28
Design#7	637450	642466	392697	16
Design#8	779966	777869	475353	16
Design#9	2780115	5315732	4468680	28
Design#10	2288833	3179391	2982620	28
Design#11	3362065	4974698	4752870	28
Design#12	885391	1162839	1608120	28
Design#13	798317	1042158	1916400	28
Design#14	587529	1286854	5309480	16

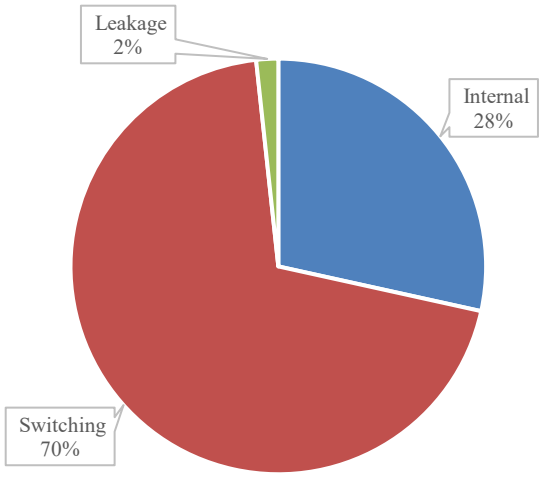


Figure 1: Static (Leakage) and dynamic (Internal + Switching) Power reparation

2.2. Wire optimization to reduce Net power

In our previous research [16], we presented a wire optimization technique for power saving on the interconnection at the physical implementation phase of an IC. At this stage, the circuit voltage and the TR are fixed by the circuit function, and we can't do anything to reduce them. The remaining parameter is the

interconnection capacitance. This represents an opportunity for significant power saving using routing transforms.

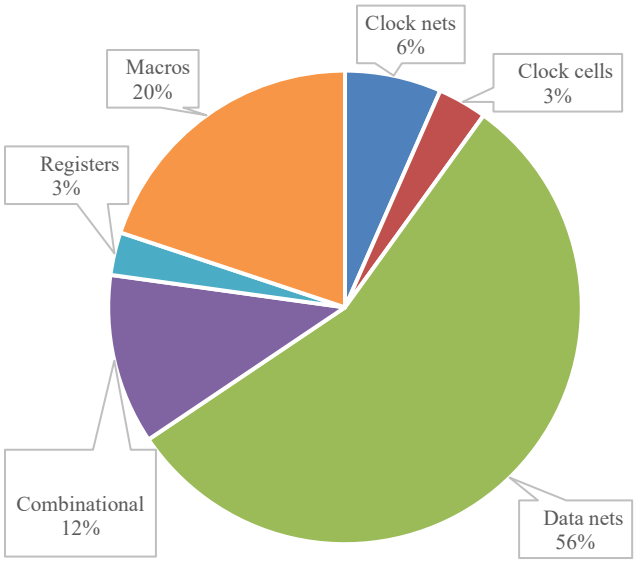


Figure 2: Multiple IC power components repartition

The complexity of capacitance variations makes it nearly impossible for the human mind to determine which combination of layers and via structures to use for a given net in order to obtain the less possible power consumption and keeping an acceptable timing and good routing. This can be achieved through layer promotion of power critical nets, coupled with a carefully set of double-spacing non-default rules (NDRs). Enabling the routing engines to efficiently trade-off timing quality of results (QoRs) and congestion.

We have used Mentor Graphics Nitro-SoC™ [22] tool and the correspondent place and route full flow [23, 24] to implement test-cases used during this study.

As a review of results achieved in our previous work, reference [16] proves the following results:

- 40% of data nets are consuming 92.6% of the total power as shown in Figure 3.
- Start having a signifying power reduction up to 20% on power consumed by data nets when the number of target nets exceeds 40% of the total number of data nets: Figure 4.
- Total power reduction percentage approaching 7 %: Figure 5.
- A good compromise between the power saving and the timing/congestion was achieved by taking the cases 40% and 30% of total data nets number as a target for power optimization.

3. Results and analysis

3.1. Algorithm enhancement for better congestion overflow

In the previous section, we have introduced our research by presenting a reminder of results achieved in previous work [16]. Experiment on one test-case shows an important total power saving gain exceeding 5% by targeting 30% of data nets for optimization.

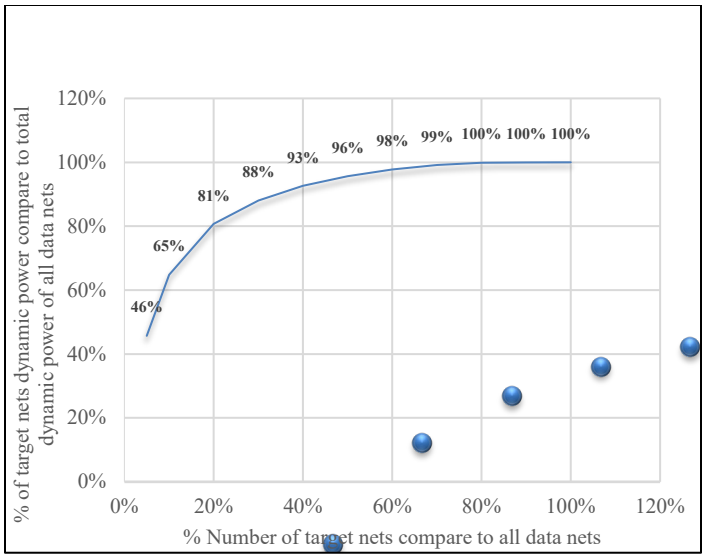


Figure 3: Data Nets power repartition

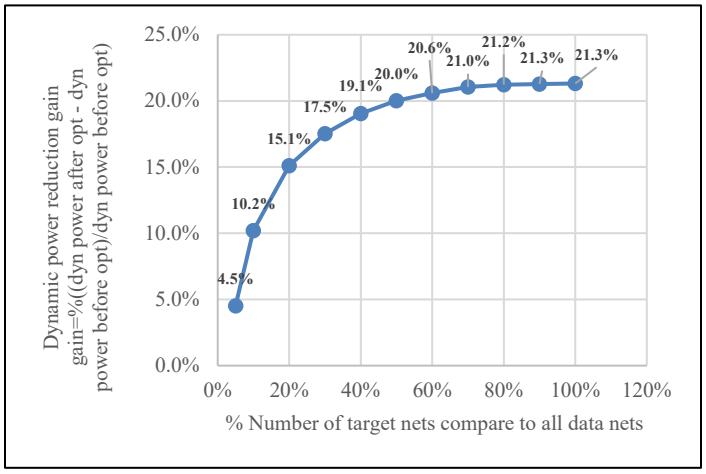


Figure 4: Dynamic power gain on data Nets

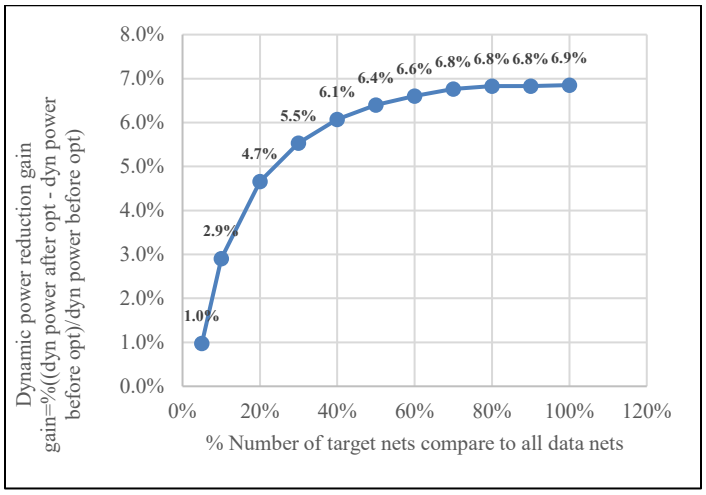


Figure 5: Total power gain on entire design

In this section, we will apply this optimal solution to multiple designs made with different technologies nodes. The goal is to see if this solution is robust enough for production usage.

For a good comparison, we are applying the wire optimization for power reduction on an optimized post-clock-tree-synthesis (post_CTS) database (db). The baseline run that produces initial post-CTS db is a full place and route flow power-driven [23, 24]. Thus, with our solution, we will be able to see the exact power gain after using existing power optimization transforms. Experiments are done on 14 test cases made with advanced technologies nodes 28 nm, 20 nm, 16 nm, and 7 nm. Their main characteristics are summarized in table 1:

Algorithm#2, shows the automatic incremental optimization flow for power reduction:

Algorithm # 2: Automatic incremental flow

1. *set designs_list = (Design#1, Design#2, Design#3, Design#4, Design#5, Design#6, Design#7, Design#8, Design#9, Design#10, Design#11, Design#12, Design#13, Design#14)*
2. ***For** design ($designs_list)*
    a. *Read baseline saved Post-CTS Database*
    b. *Report & save Initial QoR (Timing, Power, and Congestion)*
    c. *Create double spacing NDRs on all used layers for routing.*
    d. *Get 30% of data nets with highest power consumption* ***(see next section for more details: Procedure#1)***
    e. *Apply NDRs on Target_Nets list*
    f. *Run the native global route*
    g. *Run one optimization pass for Timing recovery*
    h. *Run the native global route*
    i. *Report QoR (Timing, Power, and Congestion)*
3. ***END for***

The section below describes the main procedure that gets target nets for optimization. Its objective is to get 30% of data nets having high dynamic power.

Procedure # 1: Get target nets: *30% of data nets with the highest power consumption*

1. *Filter data nets from all nets (data_nets)*
2. ***For each** data_net ($data_nets)*
    a. *Get its dynamic power and save it in a table*
    b. *Save the data_net and its dynamic power*
3. ***END for***
4. *Rank nets in order of power consumption.*
5. *Return first 30% of data_nets*

For all trials, the target nets are 30% of data nets that have the highest power consumption, while in some test cases we see that by targeting only 20% of data nets we are targeting more than 80% of the total power consumed in the interconnection. This remark conducts us to an optimal solution by targeting a dynamic list of nets consuming more than 80% of the total power.

The new procedure that gets target nets for optimization is described below:

Procedure # 2: Get target nets: nets consuming > *80% of the total power in the interconnection*

1. *Filter data nets from all nets (data_nets)*
2. *Get the total dynamic power of all $data_nets (data_nets_power)*
3. ***For each** data_net ($data_nets)*
    a. *Get its dynamic power and save it in a table*
    b. *Save the data_net and its dynamic power*
4. ***END for***
5. *Rank nets in order of power consumption.*
6. *Return the first list of nets consuming >80% of total data_nets_power*

Table 2, shows a comparison between cases performed by using procedure#1 and procedure#2. We notice an important "Overflow" reduction in almost all test-cases.

The high overflow reduction is happening on designs having a low activity such as "Design#1, Design#2, Design#3 and Design#4". The low overflow is coming especially from the important reduction of target nets from 30% to 2% for Design#1, 3% for Design#2, and 4% for Design#3&#4. For these test-cases the low target nets percentage is sufficient to have good power reductions achieving -17.5%, -16.3%, -7.6% and -7.5% respectively.

For other designs, Design#5 to Design#13, by targeting 20% instead of 30% of data nets, we achieve almost the same power gain lower congestion overflow. Finally, one test-case "Design#14" ends with the same target nets percentage of 30% and a power gain of -19%.

3.2. Power gain on overall designs

In the previous section, we notice an important reduction of the number of target nets considered for optimization with proc#2 compare to proc#1. Fewer target nets number with almost the same or better power gain is helping for a fast run time accompanied by a better rout-ability. All of that conduct to an optimal solution, which is to target the number of nets consuming more than 80% of the total net's power with a maximum of 30% of the total nets number.

Table 3, presents the dynamic and total power gain for each design using the optimal solution. An average of 5% power reduction obtained in both dynamic and total power.

Table 2: Comparison between Procedure#1 and Procedure#2.

Target nets power gain = %(nets power after optimization- nets power before optimization)/ nets power before optimization

Designs		Total number of nets	Number of target nets	% of target nets (%)	Before Optimization: Total nets power (mW)	Before Optimization: Target nets power (mW)	Before Optimization: % power of target nets (%)	After Optimization: Overflow	After Optimization: Target nets power (mW)	After Optimization: Target nets power gain (%)
Design#1	Proc#1	483171	144951	30.0%	1.0423	1.0423	100.0%	3.41	0.8752	-16.0%
	Proc#2	483171	9663	**2.0%**	1.0423	0.8578	82.3%	**1.10**	0.7079	**-17.5%**
Design#2	Proc#1	501314	150394	30.0%	1.1548	1.1548	100.0%	3.80	0.9781	-15.3%
	Proc#2	501314	15039	**3.0%**	1.1548	0.9812	85.0%	**1.69**	0.8209	**-16.3%**
Design#3	Proc#1	878226	263468	30.0%	41.998	41.8647	99.7%	0.23	38.7659	-7.4%
	Proc#2	878226	35129	**4.0%**	41.998	34.7165	82.7%	**0.15**	32.0757	**-7.6%**
Design#4	Proc#1	147116	44135	30.0%	15.435581	15.435581	100.0%	0.40	14.312023	-7.3%
	Proc#2	147116	5885	**4.0%**	15.435581	15.435581	100.0%	**0.05**	14.270356	**-7.5%**
Design#5	Proc#1	433469	130041	30.0%	102.6567	96.0036	93.5%	4.46	70.924	-26.1%
	Proc#2	433469	86694	**20.0%**	102.6567	90.8166	88.5%	**4.08**	66.463	**-26.8%**
Design#6	Proc#1	3618017	1085405	30.0%	892.1152	821.1902	92.0%	7.56	712.3573	-13.3%
	Proc#2	3618017	723603	**20.0%**	892.1152	780.1937	87.5%	**7.13**	674.5511	**-13.5%**
Design#7	Proc#1	636971	191091	30.0%	53.3274	50.7283	95.1%	2.39	39.7819	-21.6%
	Proc#2	636971	127394	**20.0%**	53.3274	48.4197	90.8%	**2.15**	37.6407	**-22.3%**
Design#8	Proc#1	770303	231091	30.0%	57.9306	55.0936	95.1%	1.75	41.678	-24.4%
	Proc#2	770303	154061	**20.0%**	57.9306	52.2459	90.2%	**1.49**	39.1292	**-25.1%**
Design#9	Proc#1	2436751	731025	30.0%	939.3153	875.8495	93.2%	4.75	726.5736	-17.0%
	Proc#2	2436751	487350	**20.0%**	939.3153	831.0609	88.5%	**4.50**	688.0895	**-17.2%**
Design#10	Proc#1	2020267	606080	30.0%	1479.458	1377.5008	93.1%	8.69	1150.6352	-16.5%
	Proc#2	2020267	404053	**20.0%**	1479.458	1297.9898	87.7%	**8.19**	1082.2901	**-16.6%**
Design#11	Proc#1	3100457	930137	30.0%	2615.8898	2411.9914	92.2%	8.11	1991.3135	-17.4%
	Proc#2	3100457	620091	**20.0%**	2615.8898	2266.4228	86.6%	**7.70**	1865.8817	**-17.7%**
Design#12	Proc#1	822186	246656	30.0%	435.9402	411.998	94.5%	7.83	346.8801	-15.8%
	Proc#2	822186	164437	**20.0%**	435.9402	394.3237	90.5%	**7.57**	332.0825	**-15.8%**
Design#13	Proc#1	745762	223729	30.0%	378.2664	338.9859	89.6%	4.86	274.266	-19.1%
	Proc#2	745762	149152	**20.0%**	378.2664	313.018	82.8%	**4.55**	253.8826	**-18.9%**
Design#14	Proc#1	466631	139989	30.0%	903.8741	790.2763	87.4%	6.30	640.3705	-19.0%
	Proc#2	466631	139989	**30.0%**	903.8741	790.2763	87.4%	6.30	640.3705	**-19.0%**

Table 3: Dynamic and total power gain

power gain = %(power after optimization- power before optimization)/ power before optimization

Designs	Before Optimization		After Optimization			
	Dynamic Power (mW)	**Total power (mW)**	**Dynamic Power (mW)**	**Total power (mW)**	**Dynamic power gain (%)**	**Total power gain (%)**
Design#1	8.0476	8.0711	7.9125	7.9359	**-2%**	**-2%**
Design#2	8.7976	8.9791	8.6429	8.8298	**-2%**	**-2%**
Design#3	149.3188	152.1707	147.3411	150.1936	**-1%**	**-1%**
Design#4	113.042473	113.31836	111.774694	112.050549	**-1%**	**-1%**
Design#5	330.4627	330.8185	307.2876	307.6482	**-7%**	**-7%**
Design#6	2658.979	2659.736	2563.1845	2563.9415	**-4%**	**-4%**
Design#7	198.2466	198.3225	187.7011	187.7769	**-5%**	**-5%**
Design#8	218.2062	218.3048	205.2356	205.334	**-6%**	**-6%**
Design#9	2018.207	2126.9193	1872.6143	1981.4807	**-7%**	**-7%**
Design#10	2964.9741	3043.4681	2748.532	2830.5278	**-7%**	**-7%**
Design#11	4141.0681	4245.7608	3735.9251	3840.9783	**-10%**	**-10%**
Design#12	894.2836	894.3742	833.1615	833.2521	**-7%**	**-7%**
Design#13	702.7373	702.8631	643.1108	643.2364	**-8%**	**-8%**
Design#14	2693.7224	2694.3472	2595.0319	2595.6581	**-4%**	**-4%**
				Average	**-5.18%**	**-5.10%**

4. Conclusion

In this paper, we present a new wire optimization technique for power reduction during IC physical implementation phase. The main outcome is the optimal choice of target nets for optimization, which is the list of power critical nets consuming more than 80% of total power in the interconnection without exceeding the number of nets of 30% of the total nets. Experiment on 14 test-cases made with advanced technologies nodes shows an important power reduction and, at the same time, keeps having good design rout-ability.

The technique leads to an important dynamic power improvement through a simple critical Nets re-routing. The power on all data Nets reduced up to 20% and the average total power reduction in all test-cases by 5%.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research supported by Mentor Graphics Corporation. We thank Dr. Hazem El Tahawy (Mentor Graphics, Managing Director MENA Region) for initiating and supporting this work.

References

[1] K. C. Janac, “Interconnect Physical Optimization,” Proceedings of the 2018 International Symposium on Physical Design - ISPD 18, 2018.

[2] D. J. Radack and J. C. Zolper, "A Future of Integrated Electronics: Moving Off the Roadmap," in Proceedings of the IEEE, vol. 96, no. 2, pp. 198-200, Feb. 2008. doi: 10.1109/JPROC.2007.911049.

[3] D. Flynn, R. Aitken, A. Gibbons, K. Shi, Low power methodology manual: for system-on-chip design, 2ed. ed., Springer, 2007, p13.

[4] I. Lee and K. Lee, “The Internet of Things (IoT): Applications, investments, and challenges for enterprises,” Business Horizons, vol. 58, no. 4, pp. 431–440, 2015.

[5] C. J. Alpert, W.-K. Chow, K. Han, A. B. Kahng, Z. Li, D. Liu, and S. Venkatesh, “Prim-Dijkstra Revisited,” Proceedings of the 2018 International Symposium on Physical Design - ISPD 18, 2018.

[6] ITRS 2015 Edition Report-Interconnect, https://www.semiconductors.org/wp-content/uploads/2018/06/6_2015-ITRS-2.0_Interconnect.pdf, 2015.

[7] J. Hu, Y. Zhou, Y. Wei, S. Quay, L. Reddy, G. Tellez, and G.-J. Nam, “Interconnect Optimization Considering Multiple Critical Paths,” Proceedings of the 2018 International Symposium on Physical Design - ISPD 18, 2018.

[8] S. Mantik, G. Posser, W.-K. Chow, Y. Ding, and W.-H. Liu, “ISPD 2018 Initial Detailed Routing Contest and Benchmarks,” Proceedings of the 2018 International Symposium on Physical Design - ISPD 18, 2018.

[9] X. Qiu and M. Marek-Sadowska, “Routing Challenges for Designs With Super High Pin Density,” IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 32, no. 9, pp. 1357–1368, 2013.

[10] L. Cherif, M. Chentouf, J. Benallal, M. Darmi, R. Elgouri and N. Hmina, "Usage and impact of multi-bit flip-flops low power methodology on physical implementation," 2018 4th International Conference on Optimization and Applications (ICOA), Mohammedia, Morocco, 2018, pp. 1-5. doi: 10.1109/ICOA.2018.8370498.

[11] M. Rahman, R. Afonso, H. Tennakoon and C. Sechen, "Design automation tools and libraries for low power digital design," 2010 IEEE Dallas Circuits and Systems Workshop, Richardson, TX, 2010, pp. 1-4.

[12] M. Chentouf, L. Cherif and Z. El Abidine Alaoui Ismaili, "Power-aware clock routing in 7nm designs," 2018 4th International Conference on Optimization and Applications (ICOA), Mohammedia, Morocco, 2018, pp. 1-6. doi: 10.1109/ICOA.2018.8370505.

[13] L. Cherif, J. Benallal, M. Darmi, M. Chentouf, R. Elgouri, and N. Hmina, "ASIC Physical Design Flow: Power Saving Opportunities on Interconnection Components," Lecture Notes in Electrical Engineering Proceedings of the 1st International Conference on Electronic Engineering and Renewable Energy, pp. 258–265, Feb. 2018.

[14] L. Cherif, M. Darmi, J. Benallal, R. Elgouri, and N. Hmina, "Last-Mile Post-Route Power Optimization in Integrated Circuit Conception," International Journal of Engineering & Technology, 7(4.16), 102-105. doi:http://dx.doi.org/10.14419/ijet.v7i4.16.21788.

[15] G. J. Y. Lin, C. B. Hsu and J. B. Kuo, "Critical-path aware power consumption optimization methodology (CAPCOM) using mixed-VTH cells for low-power SOC designs," 2014 IEEE International Symposium on Circuits and Systems (ISCAS), Melbourne VIC, 2014, pp. 1740-1743.

[16] L. Cherif, M. Chentouf, J. Benallal, M. Darmi, R. Elgouri and N. Hmina, "Layer Optimization for Power Reduction in Integrated Circuits," 2018 IEEE 5th International Congress on Information Science and Technology (CiSt), Marrakech, Morocco, 2018, pp. 625-629. doi: 10.1109/CIST.2018.8596605.

[17] J. M. Rabaey, A. P. Chandrakasan, and B. Nikolic, Digital Integrated Circuits: A Design Perspective, 2nd ed., Prentice Hall Electronics and VLSI Series, Upper Saddle River, NJ: Pearson Education, 2003.

[18] K. Y. Gary, Practical Low Power Digital Vlsi Design. Springer Verlag, 2012. doi:10.1007/978-1-4615-6065-4.

[19] M. S. Elrabaa, I. S. Abu-Khater, and M. I. Elmasry, Advanced low-power digital circuit techniques. Boston: Kluwer Academic Publishers, 1997.

[20] J-G. Cousin, D. Chillet and O. Sentieys, "Power Estimation and Optimisation for ASIPs", Submited to 1997 International Symposium on Low-Power Design, Mont. CA, Aug. 1997.

[21] E. Macii, "High Level Design and Optimization for Low Power", NATO Advance Study : Low Power in Deep Submicron Electronics, Aug 1996.

[22] Nitro-SoC™ Software Version 2018.1, December 2018.

[23] Nitro-SoC™ "User's Manual", Software Version 2018, December 2018.

[24] Nitro-SoC™"Advanced Design Flows Guide", Software Version 2018, December 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 142-150 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Into the Online Space: Outcomes for a PASS Online Pilot Across Different First Year University Subjects

Tracey Woolrych[*,1], Melissa Zaccagnini[2], Moira Stephens[3], Murray Stace[4], Rebekkha Middleton[3], Melissa Bergner[2], Sian O'sullivan[5], Reetu Verma[6]

[1]*University of Wollongong; School of Psychology; UOW, 2522, Australia*

[2] *University of Wollongong; PASS; UOW, 2522, Australia*

[3] *University of Wollongong; School of Nursing; UOW, 2522, Australia*

[4] *University of Wollongong; LTC; UOW, 2522, Australia*

[5] *University of Wollongong; Peer Learning; UOW, 2522, Australia*

[6] *University of Wollongong; School of Accounting, Economics & Finance; UOW, 2522, Australia*

ARTICLE INFO

Article history:
Received: 30 September, 2019
Accepted: 24 December, 2019
Online: 22 January, 2020

Keywords:
Peer Assisted Study Sessions
Online Learning
First year University students

ABSTRACT

Peer Assisted Study Sessions (PASS) is a student let program designed to support students transitioning into university and tackling traditionally difficult first year core subjects. The model is collaborative with student leaders facilitating activities and discussions driven by student need. Consistently research has found that students can benefit, in terms of increased grades, from attending PASS, however findings for online delivery modes are mixed. These studies have generally only compared face to face (F2F) with online modes of PASS-like programs for one subject. No study has compared different subjects from varying disciplines to investigate if the benefits of PASS online are the same for all students. PASS at UOW conducted a pilot study of synchronous online sessions, tested across three different first year university subjects. A total of 1.471 students enrolled into these subjects, with 409 attending some form of PASS. Result revealed PASS students gained significantly higher average final marks compared to students who did not attend any type of PASS, regardless off subject. However, results for PASS varied depending upon mode of delivery engaged in (F2F or online) and also with subject. Although not all differences were statistically significant, trends suggest a student/subject interaction that may vary the amount of benefit gained from PASS online formats. Possible drivers for these results are discussed as well as consideration given to cohort effects and student skills for online learning modes.

1. Introduction

As Universities strive to meet the needs of more diverse student populations, online learning options are being rolled out Australia wide. Time poor students, many of whom have families and/or work commitments, require more flexibility to complete their degrees, including a need for online-based supports [1-2]. Peer led learning assistance programs are following this trend by offering synchronous online learning sessions [3]. However it is unclear if students are benefitting from this mode of learning, especially in the first year space where students may not yet be fully independent learners [4]. Currently research into online versions of PASS-like programs have offered positive results in terms of benefits for students. However, many of these studies investigate only one subject at a time, making it unclear whether student and/or subject characteristics may influence the benefits students can obtain.

Research on PASS-like programs has shown many benefits for students including increased grades, particularly in face to face (F2F) delivery modes [5-6]. These findings seem consistent across a range of different courses [7]. However, findings reveal this

[*]Corresponding author: Tracey Woolrych, Email: woolrych@uow.edu.au

consistency is not present for online versions of PASS [8-11]. The F2F environment comes with a range of stimuli and cues which are important in facilitating learning [8]. Given the collaborative nature of Peer Assisted Study Sessions (PASS), these cues become essential for successful student/student and student/leader interactions. There is no guarantee that these can be created or re-created in the online space. Online learning modes may not be able to emulate the F2F student experience [4, 12-14]. One could argue that they should not try, given that the content and mode of delivery for different subjects are intrinsically linked. It is this that may impact on the success of online delivery. Not every subject may be suitable to an online format, with important learning-based elements lost in an online delivery. Therefore it is vital that research address the question of subject and/or student suitability for online delivery, and whether it is beneficial for student learning.

Although current peer-assisted models span a variety of discipline areas, including Engineering and IT [15], languages [8], and occupational therapy [16] few studies have compared these outcomes across multiple subjects and/or disciplines. Subjects vary in their learning outcomes, and the approaches and skill development requirements to meet these. Therefore, a one-size-fits all approach to the adaptation of F2F peer-led learning programs to online versions may not be advisable. Research comparing peer-led online formats across a range of subjects is required to unpack some of the underlying factors that may be contingent on student success. This paper seeks to address this gap by comparing different first year subjects across multiple disciplines. This allows us to explore whether PASS online formats are equally beneficial to students in different subjects and whether other factors such as student and/or subject characteristics may play a role. This is essential to the future development of online formats of PASS-like programs.

2. Peer Assisted Learning (PAL) Programs

Peer Assisted Study Sessions (PASS) was originally based on supplemental instruction (SI) developed by Dr D Martin at the University of Missouri over 30 years ago [17]. SI programs have since spread around the world with many tertiary institutions adopting them in an effort to support their students. PASS have been operating in collaboration with Australian Universities for many years. According to the Supplemental Instruction Journal (2016), SIs are informal sessions whose direction is driven by students and led by students (Leaders) who have previously done well in the subject. Modelled on SI, PASS leaders do not 'teach' as much as facilitate activities and discussions designed to help students more clearly and deeply understand the subject content [17-18]. Students also develop and practice effective study skills [19]. Collaboration is key here, a far cry from the stand and deliver techniques of many lectures: at PASS you will find leaders helping students, and students helping students [20-21].

Research into the peer-led learning space reveals many benefits for students. On average, students who attend these programs obtain higher final marks compared to those that do not attend [5-6]. A higher number of students pass the subject linked with PASS, and more students finally graduate [7]. Dancer et al., [22] also found that there were higher retention rates and few withdrawals in these subjects. Lower-achieving students gained more benefits in terms of grade increases compared to higher-achieving students [22]. The positive benefits that students can gain from PASS-like programs are clear. What is less clear is how these peer-led models of learning translate into the online space.

2.1. Findings in the online space

Increasingly today's students are facing time pressures, trying to balance work/family and study commitments. This results in students desiring more flexible study options at University. Education providers are meeting this need through the delivery of blended and online learning [23]. Online study is offered over a wide range of subjects and courses. Peer led assisted programs are following suit, offering online, mostly synchronous, study sessions for students.

Research into the benefits and success of peer lead online learning is still growing, and the literature is not of a single voice in its success. Some studies have found student benefits in the form of increased satisfaction and higher average grades e.g.:[24-28]. Students also report reduced feelings of isolation [26], enjoying the flexibility [20] and collaboration [5, 29-30] that resulted from the online sessions. Online formats of peer lead learning resulted in increased student interest and engagement [30], as well as increased understanding [29]. In particular, those subjects that required skills development in IT [15] and languages [8, 31] reported success in the online mode.

Some studies revealed no differences in average final marks between students attending F2F peer-led learning sessions, compared to online attendance [8, 26-28] suggesting, on the surface, that both modes afford students the same benefits. However other studies have found a lack of grade increases for their online cohorts [6]. These variations in results could be explained by poor student uptake for the online formats [13, 23, 32] and a lack of student participation in feedback [33]. Technical issues experienced in the online format may also result in less that optimal learning environments resulting in some students abandoning the sessions altogether [32]. Recent studies have also highlighted the need for leader training specifically for the online mode of delivery [13, 23, 32]. Terminology is also an issue in this area of research. The definitions of PASS attendance are variable which may explain some of these results. As Dawson et al., [7] point out, the distinction between peer-groups and non-peer groups is at times arbitrary with studies using cut-offs varying from 1, 3 or more classes to define group conditions. The variation and rapidly evolving nature of the technology being used may also contribute to variable findings, with online studies needing to be read in the broader context of the technology "point in time".

The lack of gains achieved in online learning formats may also be the result of insufficient student skills. These studies often involve first year university students, some of whom struggle with the self-reliant and self-driven nature of online formats [4]. They may not yet have development the necessary problem solving [34], time-management and self-management skills [35, also see 3], such as motivation and commitment, necessary to succeed [36-38]. The online learning format presents a large departure from traditional learning environments many students would have been exposed to in their prior learning. Mature age students, on the other hand, may lack the necessary digital literacy and confidence to successfully navigate the online learning environment [39-42].

Studies on blended learning models report push-back from some students against taking responsibility for their own learning [14, 43]. Students new to university may have had little experience in online learning and so lack the relevant skills required for success.

An alternative explanation may be that the different subjects attract different types of students. Therefore, there may be unseen cohort effects driving some outcomes. Few studies have compared multiple subjects from different disciplines.

Another possible explanation may be that not all subjects are suited to an online mode of delivery. Whilst peer-led learning has been conducted across a range of subjects and disciplines, it should be remembered that subjects vary in their learning outcomes, approaches and requirements, meaning a one-size fits all approach may not be appropriate. The literature is unclear on this, as some studies do not state the subject under investigation, nor the faculty within which the subject is situated. The findings do suggest that STEM (science, technology, engineering, and maths) subjects do well in the online format [7, 44], although some studies did not find any grade increases for those cohorts [30, 33]. PASS conducted for Biology, as both a STEM and nursing subject, resulted in higher grades for students who attended in both the F2F and online modes [27-28]. Other subjects that have had successful outcomes in online formats are pre-doctoral (dentist) studies [24], and Education [29]. Dennis' [45] study found no differences in average grades between F2F and online versions of PASS for physical therapy students. Overall the literature suggests that students of most subjects can benefit from PASS-like programs, however whether these benefits successfully translate into the online space is less clear. In order to identify the factors that influence PASS-online success, studies need to investigate differences or similarities in outcomes across different subjects, across various disciplines.

3. Peer Assisted Study Sessions (PASS) at University of Wollongong (UOW)

The University of Wollongong (UOW) houses the Australian National Centre for PASS. The PASS program was launched in 2002 and in the intervening years has met with much success [46]. Students report higher grades and increased confidence. Traditionally, PASS is run in parallel with a range of first year subjects to support students with their transition to University study. Attached commonly to challenging first year core subjects, PASS is non-remedial, with PASS leaders guided by student needs and concerns. Currently PASS at UOW facilitates over 40 subjects (some in multiple semesters) across all faculties and onshore campuses.

The study sessions are run by PASS leaders, who have successfully passed the subject themselves. The sessions are driven by the attending students who choose the subject content to be focused on for discussion, clarification, or practice. PASS leaders engage students with a range of learning activities and guide students to develop study strategies. Activities in the F2F sessions can include brainstorming, concept lists, filing in the missing elements of an equation, and diagram labelling. The learning environment is collaborative between students and leaders. Students are encouraged to share their thoughts and knowledge with others, sharing answers and processes, often teaching back to other students. PASS students, in this way, have a direct stake in the teaching process, enhancing their own learning experiences.

3.1. The current project – PASS Online Pilot.

At UOW the demand for PASS is high. Large first year student cohorts, as well as time and space availability, limit the number of sessions that can be offered for each subject. Students also voiced a desire for some out-of-hours sessions to accommodate work and/or family commitments. Regional students also needed a way to access PASS in the absences of a suitable leader being available on their campus. In response to this, PASS at UOW trialled an online format using Blackboard Collaborate, an online synchronous platform.

PASS is student led and student focused. The selection of an online platform required careful consideration of this. The aim of the pilot was to emulate, as much as possible, the F2F student experience. It was not the aim of the pilot to replace the usual F2F sessions, but rather supplement them, by giving students more flexibility of access. The online environment needed to align with the core values and practices of PASS.

In the spirit of the student-driven principals that lie at the heart of the PASS model, student leaders were recruited to test different synchronous online platforms, as well as host the online sessions during the pilot. Blackboard collaborate was chosen for its flexibility, ease of use, and ease of access (embedded in Moodle). The platform provides for online synchronous discussion spaces, student/student and leader/student interactions, and exchange of document-based information including power point slides.

PASS leaders tested the capabilities of the system through role-play, swapping between student and leader roles to gain a deep understanding of the student experience. These experiences were shared with other leaders as part of their training for the online system and formed the basis for written guides for use in future training.

Three large cohort subjects from three different faculties were chosen for the pilot which first ran in second session, 2017. All were first year core subjects and traditionally considered difficult and challenging by students. These included: COMM121 (Statistics for Business, Commerce); NMIH106 (Essentials of Care A, Nursing); and PSYC123 (Research Methods and Statistics, Psychology).

It was predicted that if the online version of PASS provided similar benefits to F2F sessions, then, regardless of mode, PASS students would all obtain higher average marks compared to students who did not attend PASS. It was hypothesised that there would be no statistically significant difference in average mean mark achieved between the different PASS delivery modes. Finally, it was hypothesised that, regardless of PASS mode, the more hours dedicated to PASS the greater the benefits for students in the form of higher final marks.

Subject Overviews

Two of the piloted subjects were statistics-based. COMM121 and PSYC123 are first year introductory, core, statistics subjects in the schools of Business and Psychology respectively. COMM121 deals with quantitative statistical analysis for data in

the business and economics environments. Analyses and concepts include probability; hypothesis testing; correlations; regression; and time series forecasting. Students also develop skills in the use of excel and PHStat2 programs. COMM121 is an applied subject, teaching students how to use statistics to answer questions about business success and business-based weaknesses.

PSYC123 is also a quantitative-based statistics subject designed to help students answer questions about the human condition. A core subject for Psychology and Social Science students in the faculty of Social Science, students learn concepts relating to theoretical distributions; sampling; descriptive and inferential statistical analysis (z-scores; correlation; t-tests; chi-square); hypothesis testing; and methodological considerations, including validity and reliability. Students conduct all analysis and calculations by hand with a strong emphasis on result interpretation.

Essentials of Care A (NMIH106) focuses on the skills essential for successful patient/client interactions, including communication; person-centred care; knowledge (intervention; evidence; and health science); skills practice; and documentation. The subject builds directly upon knowledge and skills gained in first session nursing subjects. Foundational science knowledge is at the core of this subject, how it relates to nursing practice and how it informs patient/client communications with other health care personnel.

4. Methods

4.1. Participants

A total of 1,471 students were enrolled across the three subjects. Students were deemed PASS participants if they attended at least on session in any mode. Those students who did not obtain a final mark or obtained a final mark of zero (0) were excluded from analysis, leaving a sample size of 1,423. Of these, 409 students attended a form of PASS across the three subjects (see table .1. below).

4.2. Procedure

PASS in the online format was run concurrently with F2F sessions and with normal subject delivery in each school, throughout the second teaching session of 2017, and was open to all enrolled students from the three subjects piloted. Students voluntarily signed up by enrolling in F2F or online PASS, or a combination of both. Students were briefed about the online pilot which outlined some of the potential benefits of the program and expected time commitments (for feedback via online survey). Students could attend as many or as few PASS as they wished, in any mode of delivery. During week 6 of session, and again in week 13, students were invited to complete a brief Survey-Monkey questionnaire to gauge their experiences with the online format. All data from the questionnaires, and end of session marks for all students in each subject, were de-identified to protect anonymity. This project had ethics approval from UOW (2017/263).

4.3. Analysis

Participants were split into four discrete groups, dependent upon the PASS mode engaged in. These were PASS F2F only (PF2F); PASS online only (PO); PASS F2F and online in combination (PO+F2F); and no PASS attendance (NP). Sample numbers in the PASS online formats were very low. Non-parametric tests yielded the same results as parametrics, therefore it is the parametric test results given here. Significant results, then, should be interpreted with caution.

5. Results

5.1. Descriptives

A total of 409 students attended PASS across all three subjects. Students in PSYC123 (m =69.71) achieved the highest grades overall, followed by NMIH106 (m =66.53) then COMM121 (m =59.12). These differences were all statistically significant ($F(2,1420)=44.69$, $p<.001$ Students who attended PASS, in any format, achieved higher average marks compared to those that did not attend. For students who attended PASS, PSYC123 students achieved the highest grades (m=73.19; CI[71.10;75.28]), followed by COMM121 students (m=70.35, CI[67.00;73.69]) and NMIH106 (m=69.39, CI[66.96;71.82]) (see table .1. below).

For COMM121, 20% of students attended only 1 session of PASS, compared to 34% of students in PSYC123, and 40% of students in NMIH106.

5.2. PASS modes

Students who attended any mode of PASS achieved higher average final marks (m = 71.24) compared to students that did not attend (m = 62.35; $t(970.462)=-9.069$, p <.001; CI [-10.82; -6.97]; see figure 1 below).

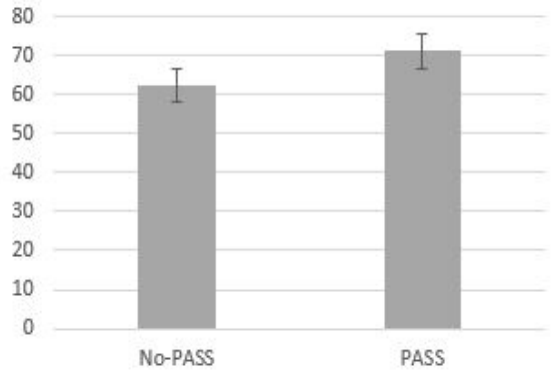


Figure 1: Mean final mark comparisons between No-PASS and PASS groups (regardless of PASS delivery mode).

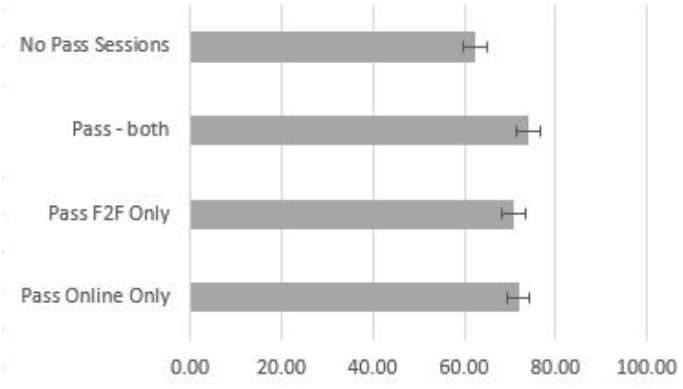


Figure 2: Mean final mark distributions comparing No-PASS and all modes of PASS delivery

Table 1: Mean final grade and mean hours of PASS attended across all PASS modes, No-Pass and for each subject.

Subject	PASS mode	n	Mean Final Grade	SD	CI	Mean hours	SD	CI
COMM121	No-PASS	411	55.81	23.13	[53.57;58.05]	-	-	
	PO	9	67.67	7.89	[61.60;73.73]	4.22	5.07	[.33;8.12]
	F2F	97	69.78	19.59	[65.83;73.73]	5.36	3.63	[4.63;6.09]
	PO+F2F	15	75.60	16.09	[66.69;84.51]	9.40	4.94	[6.66;12.14]
	Overall	***532***	*59.12*	*22.98*	*[57.16;61.07]*	*5.78*	*4.13*	[5.07;6.48]
NMIH106	No-PASS	258	65.21	12.55	[63.67;66.75]	-	-	
	PO	9	74.22	8.15	[67.96;80.49]	1.00	-	[1.00;1.00]
	F2F	104	68.72	13.63	[66.07;71.37]	3.60	2.68	[3.07;4.12]
	PO+F2F	6	73.67	14.75	[58.19;89.14]	5.33	2.16	[3.07;7.60]
	Overall	***377***	*66.53*	*12.95*	*[65.22;67.84]*	*3.49*	*2.67*	[3.00;3.97]
PSYC123	No-PASS	345	68.00	17.80	[66.12;69.89]	-	-	
	PO	15	73.00	20.23	[61.79;84.20]	4.13	3.78	[2.04;6.22]
	F2F	137	73.26	12.83	[71.09;75.42]	5.86	4.04	[5.18;6.54]
	PO+F2F	17	72.82	15.04	[65.09;80.56]	9.88	5.40	[7.11;12.66]
	Overall	***514***	*69.7082*	*16.74*	*[68.26;71.16]*	*5*	*4.36*	[5.48;6.74]
		1423						

The highest grades were achieved when students attended a combination of PASS online and F2F sessions (see figure 2 below). However, there was no statistically significant difference in grades between the different PASS delivery modes ($F(2,408)=.772$, $p=.463$) suggesting that all students benefitted from PASS, regardless of mode.

There were variations between subjects in average final marks for students who attended PASS ($F(2,408)=2.47$, $p=.086$), with PSYC123 PASS students obtaining the highest marks (see table 1 above). This difference is close to statistical significance, however, low sample sizes in some of the different PASS modes may underpower this result.

5.3. PASS modes by subject

For COMM121, PSYC123 and NMIH106 there were no statistically significant differences in average student marks across the different PASS modes. Trends indicated that, dependent upon the subject taken, the amount of mark increase varied dependent upon the mode of PASS undertaken (see figure 3 below).

For COMM121 students, the highest marks were obtained by students attending a combination of PO and F2F sessions ($m=75.60$), followed by F2F only ($m=69.78$) and then Online only ($m=67.67$). Conversely, students of NMIH106 gained higher marks when attending Online only versions of PASS ($m=74.22$) or a combination of PO and F2F ($m=73.67$), the lowest marks were achieved by students who attended F2F sessions only ($m=68.72$). PSYC123 students performed almost equally across all three delivery modes of PASS: PO (m=73.00); F2F (m=73.26); and PO+F2F (m=72.82). Whilst these differences were not statistically significant, it is interesting to note that students in different subjects gained varied benefits from the different PASS delivery modes.

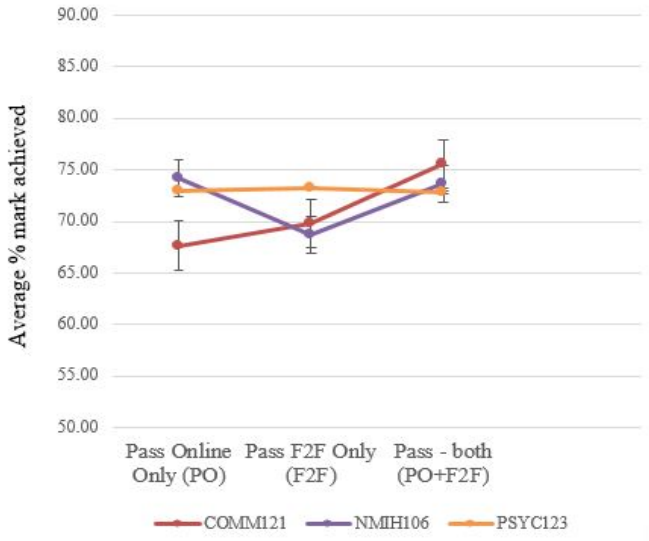


Figure 3: Comparisons of final average mark by subject and PASS delivery mode.

5.4. Correlations

To investigate the impact of hours of PASS, in its various modes, on final grade outcomes a series of correlations were conducted. Not surprisingly, there was a moderate positive, statistically significant result for number of PASS hours completed and final average marks via Pearson's Correlation Coefficient, ($r(=.226, n=409, p < .001)$. Therefore, regardless of PASS delivery

mode, the more hours completed by students, the higher their final mark.

This positive relationship held true for both NMIH106 (r=.25, n=-119, p=.006) and PSYC123 (r=.25, n = 169, p = .001). For COMM121 however the relationship was not significant (r = .16, n=121, p=.079). Looking into these correlations further, only positive relationships between hours and final marks were found for the F2F mode of PASS for NMIH106 (r=.28, n=104, p.004) and PSYC123 (r= .38, n=137, p < .001). All online modes returned non-significant results suggesting that for online versions students did not benefit from longer hours in PASS. The lack of student numbers in some of the online cohorts in this study, however, may have contributed to the non-significant results.

There were some negative relationships found, although they were non-significant. For COMM121 and PSYC123 students, hours spent in the PO+F2F mode of PASS was negatively correlated with final mark achieved, meaning that more hours did not necessarily equate to higher marks.

6. Discussion

Increasingly, institutions are offering online versions of programs to meet student demand for flexibility. Peer-led learning models have been found to benefit students in terms of final marks [5-6,22] as well as confidence [3]. However, whether these benefits can be successfully transferred to online versions is less clear. As Wang et al., [12] point out, online formats may not be able to replicate the stimuli and cues that are essential to interpersonal interactions and collaboration that can be found in F2F versions. Inconsistent results in the literature suggest this is a complex question [8] meaning other factors might be of influence. The findings from this Online PASS pilot highlight the fact that, depending upon the mode of PASS engaged in, student benefits in terms of final grades can vary. Students characteristics, which may help explain this variation, were not investigated in this current study, and should be the focus of future research.

Overall, students who attended PASS achieved significantly higher final marks compared to students who did not attend. This result seems to be driven by the F2F version of PASS, although there were no significant differences in final marks between PASS delivery modes. This is in line with previous literature that has highlighted the benefits of Peer-led learning programs [3, 24-26]. Therefore, PASS can result in positive benefits for students across a range of subjects, for the F2F version at least.

The results varied according to subject. For NMIH106 and PSYC123 there were benefits for students in the online versions of PASS, however COMM121 students seemed to struggle with this. In COMM121 in particular, numbers in the online cohorts were low which may have resulted in underpowering of the statistics used. This is not unusual according to the literature, where, despite student demand for more flexible learning options, when offered, the uptake is often low [20,32,46].

This study also found that overwhelmingly students still prefer the F2F versions of PASS. Students straight from high school, and mature age students, may be less adept at learning in the online space, preferring the more familiar class-room-like environment. This may be due to a lack of confidence, knowledge and/or skills for online formats [3-4, 34-38]. Skills development is an important issue that should be addressed for students coming into University study. Enhancing students' online learning skills will help to build their confidence in the online space, making them better able to take advantage of PASS online.

The nursing subject, NMIH106, was the only subject where students in the online only condition achieved higher marks than students in the F2F only condition. Again, this highlights the fact that different subjects are host to students of varying characteristics that can impact the level of success they achieve in online learning formats. This finding is counter to much of the literature which indicate either F2F modes result in higher grades, or that there is no difference with the online version. Although the current results were not statistically significant, it was the only trend where average final marks were higher for the online versions compared to F2F. This may be reflective of a cohort effect. Nursing students in Australia tend to be older than most other first year university students [48] with a mean age of 28.6. Salamonson et al., [48] also reported that this cohort work an average of 13.6 hours a week, although it is unclear if this is above or below the average for other first year university students. Although age and work hours were not elicited in this study it may be reasonable to suggest that the UOW Nursing cohort are not unlike that of the Salamonson et al.'s study [48]. The fact that this group performed better in the online versions of PASS perhaps speaks to more mature time-management and self-motivation skills. The availability of an online peer assisted learning option may also be particularly helpful in overcoming barriers to scheduling F2F PASS in conjunction with compulsory practical placements for this cohort and their peer leaders. These factors should be investigated further.

Interestingly, the outcomes for both COMM121 and PSYC123 were quite different, despite both being first year statistics subjects. This suggests that different student cohorts may bring with them a different set of characteristics which may influence the success of online learning. Some students may be more prepared and have better time-management and self-motivation skills that are essential for this type of learning [3-4, 34-38]. More studies comparing subject cohorts across years should be conducted in order to clarify if the effect is due to characteristics as a result of subject/degree selection or a yearly cohort effect.

The hours spent at PASS was, in most cases, positively related to final marks, which is consistent with previous research [44]. However, for both COMM121 and PSYC123 students in the PO+F2F condition a negative relationship between hours and final marks was found, although it was non-significant. Low participant numbers in the online conditions again may have contributed to the non-significant result, however it is an interesting trend. The negative relationship is difficult to explain except to say that other factors may be involved that were not the subject of this study. Both COMM121 and PSYC123 are statistics subjects, meaning students who attended both F2F and online PASS may have lacked confidence with the subject. It may also be the case that these students lacked the necessary skills to succeed, despite their PASS attendance. The role of PASS alongside other learning supports should be investigated further.

Self-selection bias is a common problem in peer-led learning program research, as it is often the most diligent students that

attend [see 7, 49-50]. Students may participate in PASS for a number of reasons, including the drive to gain higher marks, or lack of confidence in the subject. Fetner [4] made the observation that some students struggle with the self-driven nature that is often inherent in the online learning space. Despite the flexibility this mode of learning affords students, a lack of time-management and self-motivation means those students may not be maximising the benefits that are on offer. This is particularly relevant for first year university cohorts who have not yet found the most efficient and effective study methods [43]. It has also been reported that students may push-back against online learning models, believing that the responsibility for their learning lies with the educators, rather than with themselves [14]. Therefore, it is imperative that students be offered the opportunity to develop the skills that will help them succeed in the online learning space.

Overall, the current study indicates that F2F versions of PASS deliver consistent benefits to students via increased grades, across a variety of subjects. The benefits to be had in the online learning modes, however, are less clear. The current study highlights the need to further investigate student and/or subject characteristics that may impact the benefits achieved by students engaging with the online format. Other factors such as student skills, maturity, and confidence, may be important factors that impact online learning outcomes. Similarly, subject characteristics should be studied which may reveal factors regarding the suitability of different subjects for this particular learning platform.

6.1. Limitations

Low student numbers in many of the online conditions may have undermined some of the inferential statistical analyses undertaken, therefore results should be interpreted with caution. Low sample sizes in online conditions is a problem in the literature [20,32,46], where student uptake has been lacking. This problem can be overcome in the future through higher sample numbers, however, given that enrolment into online conditions is voluntary it is likely that this will continue to be a problem for future research. There are positive outcomes from this study that suggest, with the right supports and skills development, students can gain equal benefit from online versions of PASS as those obtained in F2F versions.

The high number of students who attended only 1 PASS may also have clouded results, underpowering the analysis, particularly for the correlations. Whilst some studies do not include students that have attended only 1 PASS class [7], it was decided that any student that attended any mode of PASS for at least one session be included in order to reduce the impact of low numbers in some of the online conditions.

It should also be noted that this was a Pilot study. Students at UOW had not previously had the opportunity to participate in PASS in the online format. Despite receiving specific training for the online space, PASS leaders themselves were learning along the way. Sessions at the beginning of the academic period may have been subject to technical issues, particularly in the statistics subjects where difficulty in sharing formulas and the creation of specific mathematical symbols was problematic. Therefore, the focus, in these early sessions, may have veered away from PASS activities and interactions to the implementation issues and student user troubleshooting. PASS leaders reported anecdotally that they, and their students, grew in confidence in working with the online system as the semester progressed. These lessons have now been included in the training guides and manuals for future PASS online leaders.

7. Conclusion

This study may explain some of the inconsistencies in findings for the Peer-led online learning space. Students from different subjects had differing outcomes dependent upon the type of PASS undertaken. This suggests that there may be underlying student characteristics which vary according to the subject taken, which may impact successful learning in the online space. Larger sample sizes, particularly for the online conditions, will help to clarify this, as will studies comparing results for the same subjects across years. Student characteristics such as work and/or family commitments, age, and other factors should also be investigated.

Students may also lack the necessary skills to succeed, particularly in the dimensions of self-motivation and time-management. Despite students reporting the online system easy to use, most students still showed a preference for the F2F format. This highlights the need for student skills development in the first-year space so that students can more confidently and effectively take advantage of the flexible learning that PASS online can offer.

PASS can help students achieve higher marks. Overwhelmingly students that attended PASS F2F sessions achieved significantly higher marks compared to those who did not attend. There were a high number of students that attended only 1 session of PASS and it is unclear why this might be the case. What is clear is that dedication to PASS, in the form of multiple sessions, has benefits for students.

The study provides a useful contribution to the understanding of PASS style learning environments in the online mode. Consideration of subject selection, student demographics, technology available, training of peer leaders, and digital literacy of participants are areas for further investigation. Future studies should also investigate student and subject characteristics that may impact the benefits students gain from PASS. This study, limitations considered, indicates that not all subjects may be suitable for an online PASS platform. The role of peer assisted learning in online synchronous modes and how this relates to asynchronous interactions is outside the scope of this study, however, is an area for further investigation.

Peer-led learning programs are essential in helping students transition into university study. This PASS pilot demonstrates that, for traditionally difficult first year subjects, PASS can give students the winning edge they need. In this student-driven environment, students learn not only from PASS leaders, but from other students, engaging in discussions and activities that help elucidate the more difficult concepts in a subject, and also develop the study skills necessary to succeed. Venturing into the online space for Peer-led programs comes with it challenges, however with future studies investigating student characteristics and other factors that impact online learning outcomes will help ensure the future success of PASS Online.

8. References

[1] Gillies, D. “Student perspectives on videoconferencing in teacher education at a distance” Distance Education, 29(1), 107–118, 2008.

[2] Kearsley, G. "Is online learning for everybody?" Educational Technology, 42(1), 41-44, 2002. Retrieved from http://www.jstor.org/stable/
[3] McKenzie, K., & Schweitzer, R. "Who succeeds at university? Factors predicting academic performance in first year Australian university students" Higher Education Research & Development, 20(1), 21-33, 2001.
[4] Fetner, M. "What do unsuccessful online students want us to know?" Journal of Asynchronous Learning Networks, 17(1), 13-27, 2013.
[5] Pereira, R. "UniPASS Report Semester 1 2012," Curtin University: Australia,2012. https://life.curtin.edu.au/local/docs/2012_S1_UniPASS_Report_Web_Edit.pdf
[6] Taylor, K., & Kelly, T. "Comparison of effectiveness of several peer learning programs to support 1st year chemistry" in the 2014 Australian Conference on Science and Mathematics Education, 89-90, September 2014. (formerly UniServe Science Conference).
[7] Dawson, P., van der Meer, J., Skalicky, J. & Cowley, K. "On the effectiveness of Supplemental Instruction: A systematic review of Supplemental Instruction and Peer-Assisted Study Session Literature between 2001 and 2010" Review of Educational Research, 84(4), 609-639, 2014.
[8] Yu-Chen, K., Hui-Chun, C., & Chi-Hao, H. "A learning style-based grouping collaborative learning approach to improve EFL students' performance in English courses" Journal of Educational Technology & Society, 18(2), 284-298, 2015.
[9] Brandon, D., & Hollingshead, A. "Collaborative learning and computer-supported groups" Communication Education, 48(2), 109-126, 1999.
[10] Spaniol- Mathews, P., Letourneau, L.E., & Rice, E. "The impact of online supplemental instruction on academic performance and persistence in Undergraduate STEM courses" Supplemental Instruction Journal, 2(1), 19-32, 2016.
[11] Smith, J., Wilson, S. B., Banks, J., Zhu, L., & Varma-Nelson, P. "Replicating Peer-Led Team Learning in cyberspace: Research, opportunities, and challenges" Journal of Research in Science Teaching, 51(6), 714-740, 2014.
[12] Wang, Q., Huang, C., & Quek, C.L. "Students' perspectives on the design and implementation of a blended synchronous learning environment" Australian Journal of Educational Technology, 34(1), 1-13, 2018.
[13] Beaumont, T.J., Mannion, A.P., & Shen, B.O. "From the campus to the cloud: The online Peer Assisted Learning Scheme" Journal of Peer Learning, 5(1), 1-15, 2012. http://ro.uow.edu.au/ajpl/vol5/iss1/6
[14] O'Flaherty, J., & Phillips, C. "The use of flipped classrooms in higher education: A scoping review" Internet and Higher Education, 25, 85-95, 2015.
[15] Huijser, H., & Kimmins, L. "PALS online and community building: A contradiction in terms?" in the ASCILITE 2005: 22nd Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education: Balance, Fidelity, Mobility-Maintaining the Momentum? (1). Queensland University of Technology, Teaching and Learning Support Services, 2005.
[16] Larkin, H,., & Hitch, D. "Peer Assisted Study Sessions (PASS) preparing occupational therapy undergraduates for practice education: A novel application of a proven educational intervention" Australian Occupational Therapy Journal, 55, 100-109, 2019. doi: 10.1111/1440-1630.12537
[17] Topping, K. J. "The effectiveness of peer tutoring in further and higher education: A typology and review of the literature," Higher Education, 32(3), 321-345, 1996.
[18] Longfellow, E., May, S., Burke, L., & Marks-Maran, D. "'They had a way of helping that actually helped': a case study of a peer-assisted learning scheme" Teaching in Higher Education, 13(1), 93-105, 2008. doi: 10.1080/13562510701794118
[19] Capstick, S., & Fleming, H. "Peer Assisted Learning in an Undergraduate Hospitality Course: Second Year Students Supporting First Year Students in Group Learning." Journal of Hospitality, Leisure, Sport and Tourism Education 1(1), 69–75, 2002.
[20] Brown, K., Nairn, K., van der Meer, J., & Scott, C. "'We were told we're not teachers….it gets difficult to draw the line': Negotiating roles in Peer-Assisted Study Sessions (PASS)" Mentoring & Tutoring: Partnership in Learning, 22(2), 146-161, doi: 10.1080/13611267.2014.902559
[21] Cusick, A., Camer, D., Stamenkovic, A., & Zaccagnini, M. "Peer Assisted Study Sessions for research trainees" Journal of Peer Learning, 8(14),18-33, 2015. https://ro.uow.edu.au/ajpl/vol8/iss1/4
[22] Dancer, D., Morrison, K., & Tarr, G. "Measuring the effects of peer learning on students' academic achievement in first-year business statistics" Studies in Higher Education, 40(10), 1808-1828, 2015.
[23] Lim, C., Anderson, F., & Mortimer, R. "PASSport to the Cloud- Results of a Peer-Assisted Study Sessions (PASS) Online Pilot Program" Student Success, 7(2), 59-64, 2016.
[24] Carter-Hanson, C., & Gadbury-Amyot, C. "Implementing Supplemental Instruction online to create success in high-stakes coursework for pre-doctoral dental students" Supplemental Instruction Journal, 2(1), 53-75, 2016.
[25] Court, S., & Molesworth, M. "Course-specific learning in peer assisted learning schemes: A case study of creative media production courses" Research in Post-compulsory Education, 13, 123-134, 2008.
[26] Evans, M.J., & Moore, J.S. "Peet tutoring with the aid of the internet" British Jouranl of Educational Technology, 44(1), 144-155, 2013. doi:10.1111/j.1467-8535.2011.01280.x
[27] Finlay, K., & Mitchell, M. Z. "Evaluating the effectiveness of online vs. face-to-face Supplemental Instruction in introductory biology courses" unpublished, University of Regina, 2017.
[28] Hizer, S. E., Schultz, P., W & Bray, R. "Supplemental Instruction Online: As Effective as the Traditional Face-to-face model?" Journal of Science, Education & Technology, 26(1), 100-115, 2017.
[29] Bone, J., & Edwards, S. "Connecting in rhizomic spaces: Peer-assisted learning (PAL) and e-learning in teacher education" Australian Journal of Adult Learning, 55(1), 54-74, 2015.
[30] Edwards, S., & Bone, J. "Integrating peer assisted learning and eLearning: Using innovative pedagogies to support learning and teaching in higher education settings" Australian Journal of Teacher Education, 37(5), 1-12, 2012.
[31] Dekhinet, R., Topping, K., Duran, D., & Blanch, S. "Let me learn with my peers online! Foreign language learning through reciprocal peer teaching" Innovate: Journal of Online Education, 4(3), 2, 2008.
[32] Watts, H., Malliris, M., & Billingham, O. "Online Peer Assisted Learning: Reporting on practice" Journal of Peer Learning, 8(8), 85-104, 2015.
[33] Rourke, L., & Anderson, T. "Using peer teams to lead online discussions" Journal of Interactive Media in Education, 2002, ART.1, 2002.
[34] McPherson, M., & Nunes, M.B. "The failure of a virtual social space (VSS) designed to create a learning commuinity: Lessons learned" British Journal of Educational Technology, 35, 305-321, 2004.
[35] Lew, M.D.N., Alwis, W.A.M., & Schmidt, H.G. "Accuracy of students' self-assessment and their beliefs about its utility" Assessment & Evaluation in Higher Education, 35(2), 135-156, 2010. doi: 10.1080/02602930802687737
[36] Brouwer, J., Jansen, E., Hofman, A., & Flache, A. "Early tracking or finally leaving? Determinants of early study success in first-year university students" Research in Post-Compulsory Education, 21(4), 376-393, 2016. doi: 10.1080/13596748.2016.1226584
[37] MacCaskill, A., & Denovan, A. "Developing autonomous learning in first year university students using perspectives from positive psychology" Studies in Higher Education, 38(1), 124-142, 2013. doi:10.1080/03075079.2011.566325
[38] Mason, H.D. "Grit and academic perofrmance among first-year university students: A brief report" Journal of Psychology in Africa, 28(1), 66-68, 2018. doi:10.1080/14330237.2017.1409478
[39] Hryciw, D.H., Tangalakis, K., Supple, B., & Best, G. "Evaluation of a peer mentoring program for a mature cohort of first-year undergaduate paramedic students" Adv Physiol Educ, 37, 80-84, 2013, doi:10.1152/advan.00129.2012
[40] Cullity, M. "Challenges in understanding and assisting mature-age students who participate in alternative entry programs" Australian Journal of Adult Learning, 46(2), 175-201, 2006
[41] Spedding, J., Hawkes, A.J., & Burgess, M. "Peer Assisted Study Sessions and student performance: The role of academic engagement, student identity, and statistics self-efficacy" Psychology Learning & Teaching, 16(1), 144-163, 2017. doi: 10.1177/1475725716687166
[42] Seiber, V. "Diagnostic online assessment of basic IT skills in 1st-year undergraduates in the Medical Sciences Division, University of Oxford" British Journal of Educational Technology, 40(2), 215-226, 2009. doi:10.1111/j.1467-8535.2008.00926.x
[43] Smart, K.L., & Cappel, J.J. "Students' perceptions of online learning: A comparative study" Journal of Information Technology Education, 5, 201-219, 2006.
[44] Ng, R., Kaur, A., Sheikh Mohamed, S. F., Latif, L. A., & Bahroom, R. "E-mathematics: pre-instructional and supplement instruction and their impact of student's online participation and final exam score" Asian Association of Open Universities Journal, 4(1), 27-36, 2009.
[45] Dennis, J. K. "Problem-based learning in online vs. face-to-face environments" Education for Health, 16(2), 198-209, 2003.
[46] Zaccagnini, M., & Verenikina, I. "Peer assisted study sessions for postgraduate international students in Australia" Journal of Peer Learning, 6, 86-102, 2013.http://ro.uow.edu.au/ajpl/vol6/iss1/8
[47] Nikolic, S. & Nicholls, B. "Exploring Student Interest of Online Peer Assisted Learning Using Mixed-Reality Technology" Advances in Intelligent Systems and Computing, 715, 48-54, 2017.

[48] Salamonson, Y., Ramjan, L., Lombrado, L., Lanswer, L.D., Fernandez, R., & Griffiths, R. "Diversity and demographic heterogeneity of Australian nursing students: a closer look" International Nursing Review, 59(1), 59-65, 2011.
[49] Kochenour, E.OD.S., Jolley, J.G., Kaup, D.L., Roach, P.K.D., & Wenzler, L.A. "Supplemental Instruction: An effective component of student affairs programming" Journal of College Student Development, 38(6), 577-586, 1997.
[50] MCarthy, J. "Blended learning environments: Using social networking sites to enhance the first year epxerience" Australiasian Journal of Educational Technology, 26(6), 729-740, 2010.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 147-157 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Coordination between Heterogeneous Models Using a Meta-model Composition Approach

Naima Essadi*, Adil Anwar

SIWEB, E3S, EMI, Mohammed V University in Rabat, Rabat, Morocco

ARTICLE INFO

Article history:
Received: 17 September, 2019
Accepted: 26 October, 2019
Online: 25 November, 2019

Keywords:
Coordination
Transformation
Composition
DSML
Heterogeneity
Model
Meta-model
Offered Interface
Required Interface
Realization
Referencing
Relationship
Bridge Pattern

ABSTRACT

The technological advance as well as needs of human beings made that systems became more and more complex. In contrast, the use and creation of new modelling languages became simple and no more reserved for a handful of language experts. Consequently, many new practices of systems implementation emerged, among them, the use of different domain specific modelling languages (DSMLs) to represent the same system. Indeed, complex systems are composed by many components sometimes belonging to various domains. Thus, many teams of experts collaborate to develop such systems. Moreover, teams tend to use different DSMLs to design their concerns. This new practice generates an accidental heterogeneity due to production of various heterogeneous models representing a same system. However, those heterogeneous models need absolutely to be coordinated to facilitate communication between stakeholders and of course to ease implementation and validation of systems. This paper proposes a composition interface–based approach to coordinate and integrate heterogeneous DSMLs in order to coordinate their models. The proposed composition interface is defined according to Bridge Design Pattern. To illustrate this approach two DSMLs are used: An Indoor Service Transport Modelling Language and an Internet of Things Modelling Language.

1. Introduction

Model Driven Engineering (MDE) has as main goal a deep separation between business and technological concerns. It makes models in center of software and systems engineering. Hence, it provides concepts and tools assuring to bridge the gap between problem-level abstractions and implementation-level concepts.

However, the emergence complexity of nowadays systems raises numerous new conception and implementation challenges. Indeed, these systems involve many different domains. Consequently, many teams of experts contribute to implement a same system.

Moreover, the need and the large use of MDE as well as the popularization of software language engineering (SLE) induce the new practice of using specific hand maid modelling languages baptized Domain Specific Modelling Languages (DSMLs)

Recently, the use of DSMLs rather than Unified Modelling ones increases due to many reasons. Actually, Domain Specific Modeling languages are expressive and allow a concise and accurate specification in addition of a high level of abstraction [1, 2]. These highlights have a direct influence upon productivity and costs. Indeed, the use of domains vocabulary means less time and effort in modelling phase. It also means a high quality of communication between stakeholders thereby decreasing error rate and modelling iterations.

Actually, every domain has specific vocabulary, concepts and paradigm. This difference implies the use of a different DSML for every different domain involved in the same system.

Consequently, the use of multiple DSMLs to design systems induces an accidental heterogeneity. Actually, as result we get many heterogeneous models for a same system as illustrated by Figure 1. The diagram of Figure 2 summarises problematic causes and effect.

*Corresponding Author: Naima Essadi, SIWEB, E3S, EMI, Mohammed V University in Rabat, Rabat, Morocco, Email: naimaessadi@gmail.com

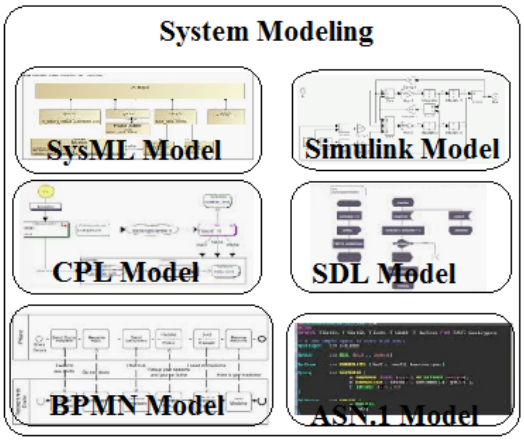


Figure 1: Heterogeneous Models for a same system

These models need inevitably to be coordinated and integrated to get a whole view of systems and, by the way to ease implementation and validation. Coordination between these models is also motivated by the need of getting relevant information scattered in different models as well as the concern to maintain the consistency of systems in case of evolution.

Many leads had been explored by current researches to resolve this issue [2, 3, 4, 5]. Specifically, this work fits in researches operating at meta-level [6, 7, 8, 9]. Actually, we choose to coordinate DSMLs used to elaborate heterogeneous models rather than coordinating models themselves. The need of a generic, reused and less error prone approach motivated this choice.

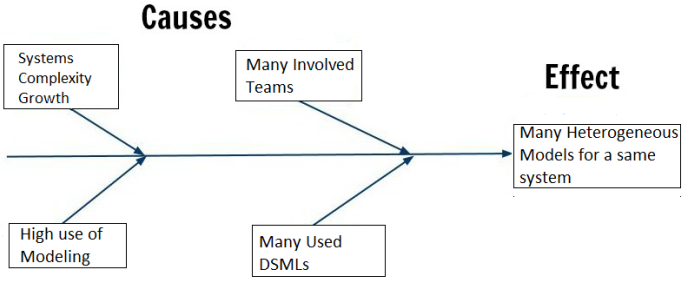


Figure 2: Causes effects diagram of current problematic

Adopting interface concept to coordinate the use of multiple DSMLs has been proposed by few interesting works [10, 11, 8, 12, 13]. Some of this works didn't give clear specification of how to define interfaces for languages, others didn't give any concreate process or tool to fulfil the proposed approach. Furthermore, those approaches didn't explain how to deal with pre-existent models.

Therefore, in this paper, we introduce a composition interface-based approach to coordinate heterogeneous DSMLs. This approach involves the application of a process in two steps: Coordination Meta-Model (CMM) elaboration and Models coordination. While, the first step aims to define interfaces for involved DSMLs and their composition to get a coordination meta-model (CMM), the second one involves the import of heterogeneous models in a model conforming to CMM. This import enables coordination between pre-elaborated heterogeneous models.

Our approach is in fact a structural interface-based composition of pre-elaborated models. It first gives a global view of system on hand and secondly affords the possibility to exchange important information between heterogeneous models as well as insuring their interoperability.

The remainder of this paper is organized as follows: we begin by giving an overview for DSMLs specification as well as heterogeneity in section 2. In Section 3, a motivating example is presented to advocate the need of resolution. Then, in following section, we introduce interface-oriented design basics for languages as well as Bridge Pattern Design. Section 5 illustrates the applicability of the approach using the motivating example where heterogeneous models are coordinated. Section 6 presents related works and finally, section 7 concludes this paper.

2. Heterogeneous DSMLs

2.1. DSMLs Overview

Domain Specific Modeling Languages aims to abstract problems and solution using domains vocabularies and concepts instead of generic and unified objects and shapes.

Contrary to a General-Purpose Modeling Language like UML, a DSML is tailored to a specific specialized domain. It allows stakeholders to contribute in modeling using notations close to their knowledge of respective domains [1]. DSMLs could be visual such as SDL, IFML, BPMN or textual like Data Modeling Languages: Abstract Syntax Notation (ASN.1) [14, 15], YANG [16, 17] and others.

DSMLs are first of all languages and consequently adhere to languages norms. Commonly, languages are specified using grammars with the famous Context-Free Grammars EBNF (extended Backus-Naur Form) [18]. Indeed, many other forms of specification exist [10] such as: Attribute Grammars [19], Graph Grammars [20], UML Profiles [21] and Meta-modeling [22]. Specifically, this work emphasizes meta-modeling specification.

2.2. DSMLs Specification

DSMLs are defined by two manners: white box specification and black box specification. The white box provides all necessary information about the language: the exhaustive list of including concepts with their allowed inter-relations as well as their concrete representation and optionally semantics giving the related meaning.

A white box classical specification reflects the language complexity. It is for example used to elaborate a compiler for the language or either to integrate languages to improve their expressiveness. Many recent researches use this specification to reuse existing DSMLs. Table 1 summarizes elements of white box specification.

However, a DSML black box specification is optional and aims to hide complexity and irrelevant information about a language. It could be defined by providing offered and required interfaces for a language. Table 2 describes offered and required interfaces.

Table 1: DSML White Box Specification

Parts	Relevance	Description
Abstract Syntax	Required	Provides concepts of a language with their exact rules of combinations [23].
Concrete Syntax	Required	Provides textual or graphical notation and symbols of language concepts [10]. A language could have one or more concrete syntax [23].
Semantics	Optional	Provides meaning and subjective understanding of a language expressions and concepts combination [10, 23]. Many types of semantics exist: Translational, Operational Semantic, Denotational and Extensional [23].
Semantics Mapping	Optional	Provides correspondences and relations between Semantic and Abstract syntax elements [23].
Syntactic Mapping	Optional	Assigns syntactic constructs to the abstract syntax [24]

The definition of interfaces for a DSML allows dealing with the language without knowing all detailed information and specification about it. The black box specification gives just restricted and needed elements. Actually, the definition of a DSML as a black box exposing specific interfaces allows language reuse in different contexts. Hence, a language could be substituted by another as long as it respects the same interface contract.

Table 2: DSML Black Box Specification

Parts	Relevance	Description
Offered Interface	Optional	Exposes elements to be used or referenced by other languages.
Required Interface	Optional	Represents elements and references needed by the language from other languages.

2.3. Heterogeneous DSMLs

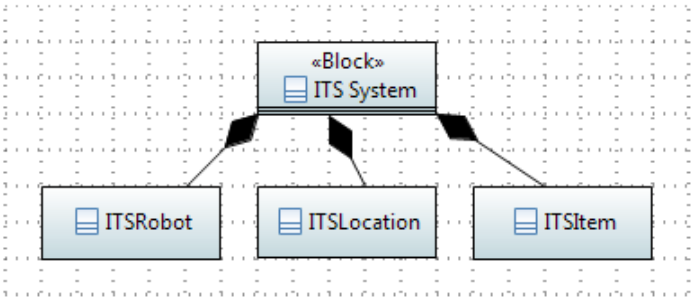
We qualify, as heterogeneous, elements with different nature. In MDE, the heterogeneity of models is usual due to various domains, tools, concepts and paradigms of modelling. Many types of heterogeneity have been identified [2, 3, 25]. Nevertheless, we notice three levels of heterogeneity between models:

- Heterogeneity of points of views: includes models with different purposes and scopes. For example a model representing static point of view is considered as heterogeneous with a model describing a dynamic point of view or a requirements model.
- Heterogeneity of meta-model (language): means that models are elaborated with different DSMLs. We can also say that they are conform to different meta-models, e.g., language heterogeneity between a BPMN model and an SDL model.
- Heterogeneity of meta-meta-model means that models are elaborated with different DSMLs, and those DSMLs themselves are specified using different meta-models. An example of that form of heterogeneity could be a CPL model elaborated using CPL DSML conforming to GOPPRR [26] meta-meta-model, the second model could be an SDL model elaborated using SDL DSML which is specified with Ecore meta-model.

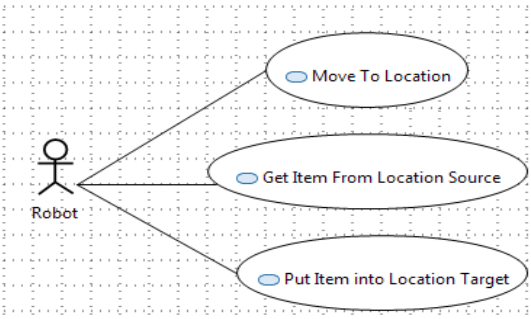
In this work we are concerned by heterogeneity of DSML which is the second level of heterogeneity. The first and third levels of heterogeneity are beyond the scope of this paper.

3. Motivating Example

In this paper we illustrate our approach using an Indoor Transport Service System (ITS). The ITS system is composed by Robots, Locations and Items as described by diagram in left of Figure 3. Robots execute tasks within a defined world [27]. It transports Items from source to target locations. This kind of systems is of course very useful and is more and more used for different purposes: In hospitals for providing patient rooms with medications and medical supplies, in restaurants to deliver ordered meals to customers, etc.



(a)



(b)

Figure 3: ITS System Main Elements and Features

However, the ITS system needs to be supervised. System users need to be informed about Robot's location and activity while accomplishing their transportation service. They also need to get informed about item's location in real time. A good solution for Robot and Items supervision is to consider them as connected objects.

This means that we need to compose our ITS system with an Internet of Things system.

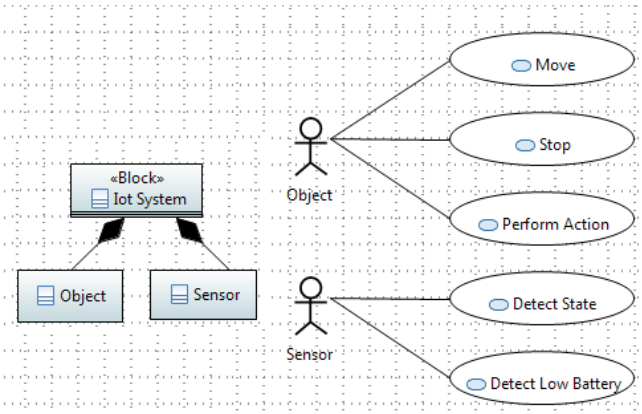


Figure 4: IoT System Main Elements and Main Features

An IoT system as described by left part of Figure 4 is composed by Objet and Sensor. Many kinds of sensors exist: sensors to detect state change, location or battery thresholds. A sensor is related to an object to detect events over them. The right part of Figure 4 exposes object and sensor use cases. Sensor detects event over object that performs actions. Consequently, in this case study, we will deal with two heterogeneous domains: ITS domain and IoT Domain. Subsequently, the composition of these two systems will involve two experts' teams: ITS team and IoT team. Every team needs a dedicated DSML for modelling its concerns, ITS DSML and IoT DSML for instance.

3.1. ITS DSML

We have elaborated a specific DSML for the ITS System. This DSML uses domain vocabularies and concepts. In this section, we will give an excerpt of the white box specification of this language. The abstract syntax is illustrated in Figure 2 and described by semantics column of table 3. The graphical representation of DSML concepts and semantics are given in Table 3.

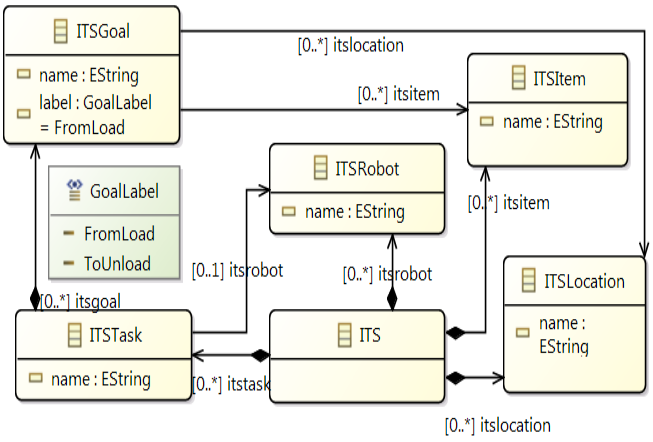
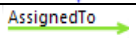


Figure 5: ITS DSML Abstract Syntax

As we can see in Figure 5, ITS DSML defines following new concepts: ITSRobot, ITSTask, ITSGoal, ITSItem and ITSLocation. ITSRobot has to achieve ITSTasks. Every ITSTask is composed by two ITSGoals: an ITSGoal to get an ITSItem from an ITSLocation source and an ITSGoal to put an ITSItem into an ITSLocation target.

Table 3: ITS DSML Concrete Syntax and Semantics

Abstract Concepts	Graphical Representation	Description
ITSRobot		Concept responsible for indoor transportation.
ITSGoal		Concept describing a mission to be executed by robot. It includes an item as well as two locations source and target.
ITSTask		Concept regrouping a list of goals to be executed by a single robot.
ITSItem		Concept representing elements transported by robots.
ITSLocation		Concept representing site of departure and end points for robots missions.
Relation ITSTask-ITSRobot		Concept relating an ITSTask to an ITSRobot
Relation ITSGoal-Location		Concept relating a goal to Location concept.
Relation ITSGoal-Item		Concept relating a goal to Item concept

We have used Obeo Designer [28] to create and design a new DSML for ITS domain. Obeo Designer is an eclipse modeling tool that enables the creation of new DSMLs and their editors. It is based on the frameworks EMF, GEF [29] and GMF [30].

Figure 6 illustrates the elaborated ITS language editor. It illustrates also a model of this DSML produced using the graphical editor. In this model the instance "Hospital-Robot" of ITSRobot has the ITSTask "Hospital_Task" as mission to achieve. This task is composed by two ITSGoals: a goal "Source" to load ITSItem "oxygenBottle" from ITSLocation "Storage" and a second ITSGoal "Target" to transport "oxygenBottle" to "PatientRoom".

A second model of this DSML is given in Figure 7. This model concerns food delivery where an "FooDeliveryRobot" has to accomplish the ITSTask "Pizza Delivery Task" of delivering ordered food "Pizza" to a specific ITSLocation "CustomerAddress".

3.2. IoT DSML

We have also elaborated a second DSML to cover internet of things domain [25]. This DSML introduces several basic concepts of IoT domain such as: Object, Sensor, Event and Action concepts. Sensors are associated to real word objects to detect events over them. Subsequently, actions are triggered according to detected events. The abstract syntax of IoT DSML is illustrated in Figure 8.

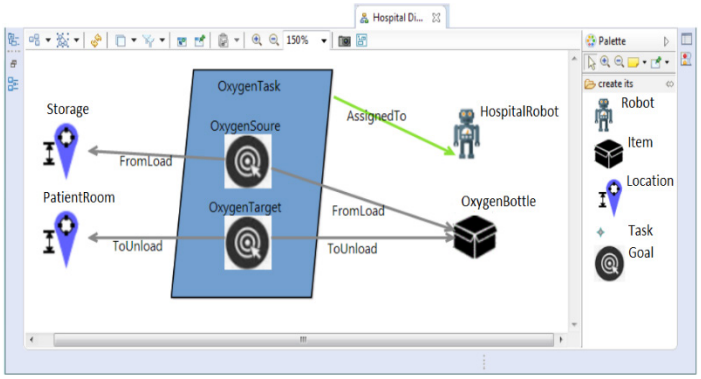


Figure 6: ITS Model for Hospital Use

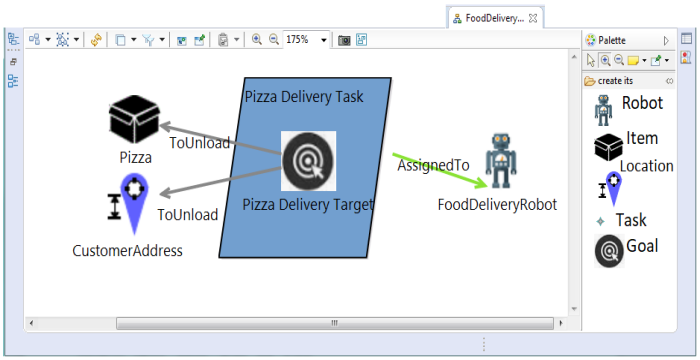


Figure 7: ITS Model for Food Delivery

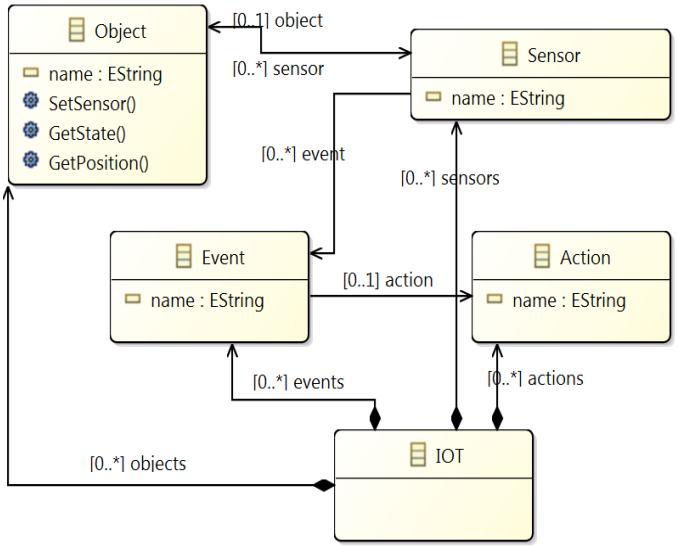


Figure 8: IoT DSML Abstract Syntax

The graphical representation and semantics of this DSML are described in Table 4.

An IoT model is illustrated in Figure 9. In this example, the instance of Object "Device" has two sensors "SpeedSensor" and "LocationSensor". The "SpeedSensor" detects event "Moving" over "Device" which trigger the action "SendSms". However, the sensor "BatterySensor" detects the event "LowBattery" and then triggers the action "SendSms".

Table 4: IoT DSML Concrete Syntax and Semantics

Abstract Concepts	Graphical Representation	Description
Object		Concept representing connected objects.
Sensor		Concept sensor responsible of detecting events over connected objects.
Event		Concept event detected by sensors over objects.
Action		Concept describing action done when a sensor detects an event or an environment change.
Relation Event-Action		Concept relating an Event to a triggered Action.
Relation Sensor-Event		Concept relating a Sensor to detected Event.

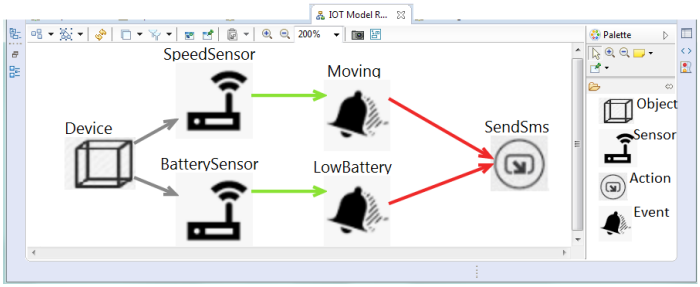


Figure 9: IoT DSML Model and Editor

A second model for IoT is given by Figure 10. The instance "Building" has "PositionSensor" as sensor. The action "SendSms" is triggered when the event "PositionChanged" is detected on "Building" Object.

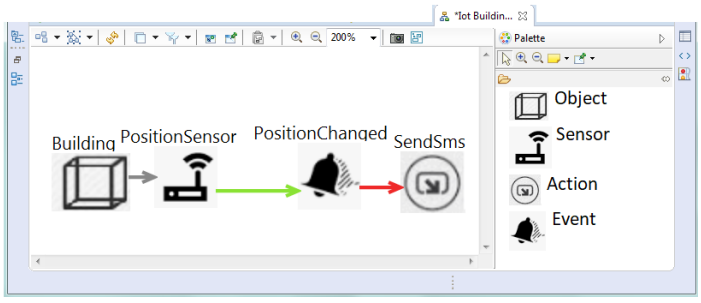


Figure 10: IoT DSML Second Model

3.3. ITS and IoT DSMLs Coordination Motivations

Both models given in Figure 6 and Figure 7 need to be related to model in Figure 9 for the following purposes: (i) To get activity and position information of "Hospital_Robot" and "FooDeliveryRobot". (ii) To get items' position and situation of orders. However, coordinating ITS and IoT DSMLs is more

beneficial than coordinating their models. As, the coordination at language level holds to all its models [10]. It actually guarantees less elaboration time, less effort and less error prone. The coordination is done once instead of for every model.

Moreover, this operation is validated once for all which decreases bugs occurrences and coordination iterations. However, to achieve coordination we need to answer the following questions:

How can we coordinate the two DSMLs to get a global view of this system?

How can we coordinate the two DSMLs to achieve the whole system goals described by (i) and (ii)?

How can we represent and describe coordination between the two DSMLs?

4. DMLs Coordination

The recent design practices, for instance the use of different DSMLs to describe domains concerns, induce an accidental heterogeneity. Indeed, system designers from different teams produce heterogeneous models to describe a single system. These models are heterogeneous as they are expressed in different DSMLs. Consequently, coordination between involved DSMLs is needed to facilitate systems elaboration.

Actually, coordination could be seen as a combination between languages [10] to work together. It aims to join DSMLs's separate concepts to achieve a common goal. This join could however be structural to get hole static and global view of systems and either to get needed information and data among systems. On the other hand, the join could also be behavioral to achieve simulation to validate systems features and execution in early implementation stages.

Furthermore, coordination is considered as a less invasive form of integration [31]. It is done posteriori to get global analysis or global simulation of a system [32]. For same purposes, many interesting works [33, 6, 7] invoked instead globalization concept. This term is used as analogy with world globalization relationships between countries to regulate interchange, interaction and communication [6].

4.1. Interface-Based Coordination Approach

We propose to consider heterogeneous meta-models as component Figure 11. We relate them using interfaces considered as coordination points. Interfaces must be defined and related internally to their own meta-models and externally to other heterogeneous meta-models using coordination relationships.

4.2. Interface Description

DSMLs interfaces are part of black box modeling languages specification. It aims to define concepts exposing possible coordination join points and relationships to be established between languages.

Indeed, the use of an offered interface of a DSML1 by a DSML2 supposes that DSML1 exposes an offered interface and in the other side the DSML2 in its turn exposes a required interface.

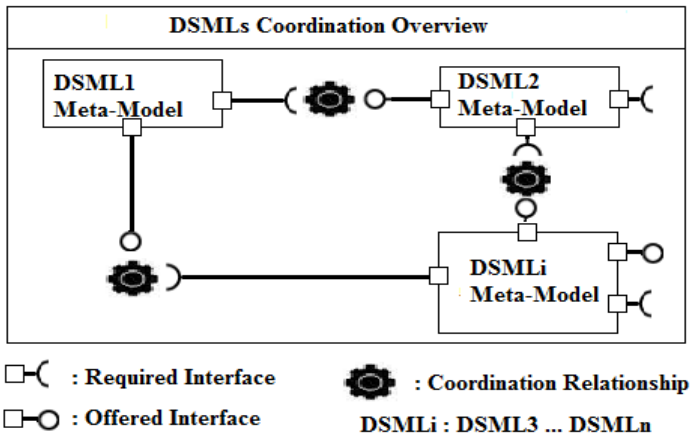


Figure 11: Heterogeneous DSMLs Coordination Overview

Our idea is described in Figure 11, where we define interfaces over DSMLs abstract syntax (meta-model). These interfaces are exposing concepts to other DSMLs and hiding details about exact structures and implementations.

In coordination, a DSML could be passive or active. The active DSML is the one exposing a required interface and thus uses interfaces offered by other DSMLs qualified as passive.

In this paper we consider external coordination between involved DSMLs. The external coordination has been defined in our previous work. It is done externally and is assured by a specific framework or workbench.

Many recent works [11, 8] discussed definition of DSMLs interfaces for various needs. However, they gave different proposition for defining it. In this work, we propose to create interfaces according to the Bridge Design Pattern [34].

4.3. Bridge Design Pattern Overview

Bridge Design Pattern has been first introduced by E.Gamma et al [34]. This design pattern aims to decouple an abstraction from its implementation so that the two can vary independently [34].

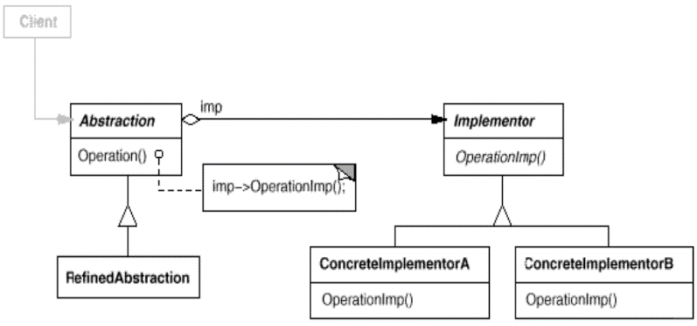
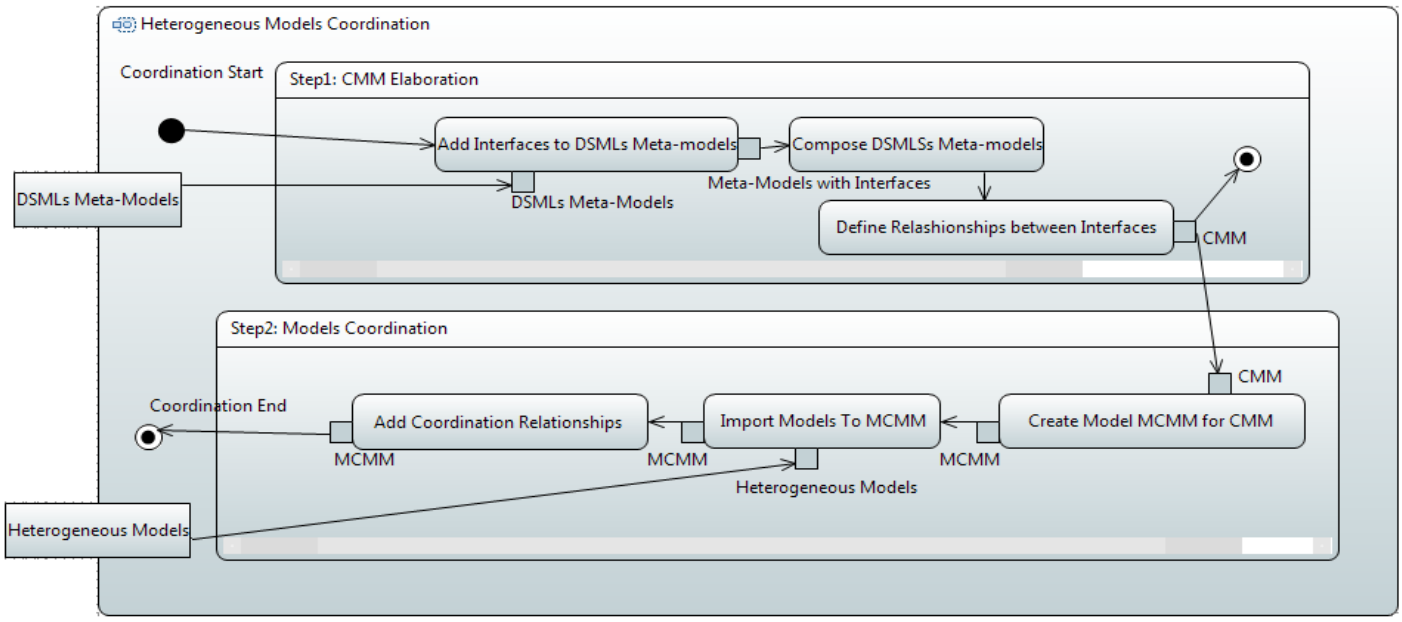


Figure 12: Bridge Design Pattern Concept [34]

The main idea of this design pattern illustrated in Figure 12 consists in the definition of an abstract class in the first part ("Abstraction" of Figure 12). This abstract class is inherited by the class requiring a specific implementation ("RefinedAbstraction" of Figure 12).



The defined Abstract class ("Abstraction") references an interface defined in the other side (see "Implementor" in Figure 12) which is implemented by classes of this same side (see "ConcreteImplementorA", "ConcreteImplementorB" in Figure 12). In case of DSMLs, we propose to define interfaces according to the Bridge Design Pattern.

Consequently, the abstraction part of the Bridge is the required interface and is defined by the first DSML while the second DSML defines the second part of the Bridge which is the interface to be referenced. This interface is exactly the offered interface of the second DSML. This means that an abstract meta-class must be defined in the first DSML. This meta-class is the super meta-class for meta-classes requiring coordination. While this meta-class represents a required interface for the first DSML, an interface is defined in the second DSML to represent its offered interface. This interface is implemented by elements of the second DSML.

The use of this pattern gives a generalized nature to coordination. Indeed, one of the two coordinated DSML could be easily replaced by another DSML as well as interfaces contracts are respected. In the proposed coordination process, a DSML could be coordinated to as many as needed DSMLs. It can be active or passive or the two at the same time.

4.4. Coordination Process Description

We introduced in Figure 11 a high-level overview of our approach. In this paragraph we propose to use this approach as a part of a coordination process composed by two main steps: CMM Elaboration and Models Coordination Figure 13.

"CMM Elaboration", represented by Fig. 13 and Fig. 14, is done at DSML level. This step is done in three stages. We begin by adding interfaces to involved DSMLs's meta-models and then we compose them to be able to define coordination relationships between their interfaces.

Interfaces are added using a transformation, while composition is done by creating a new element to be the root element of the CMM. The composition consists in including DSMLs's meta-models root elements in this new element.

We assume that in a meta-model every meta-class must have at most a parent. At least one meta-class doesn't have a parent. This meta-class is considered as the root element of a meta-model.

Subsequently, the result of this first step is coordination meta-model (CMM) that has a root element and coordinated meta-models as leaves

The second and final step is "Models coordination" illustrated by Figure 13 and Figure 15. The second step in this process enables us to achieve coordination.

We start by creating a new model MCMM conforms to CMM. The former will be composed by all input models and their mutual relationships. We first import input heterogeneous models then we relate input models interfaces using relationships defined in step 1.

The import operation is in fact a transformation that aims to transform input models to models that are conform to CMM. The result of this step is a model representing a global view of the whole system as well as its inter-relationships.

5. Application: Connected Indoor Transport Service System

To demonstrate the applicability of the proposed approach, we use the motivating example described in section 3. Actually, in this example we have two heterogeneous DSMLs to coordinate in order to get activity, position and battery threshold of ITSRobot. This coordination aims also to get position of Item and Room elements.

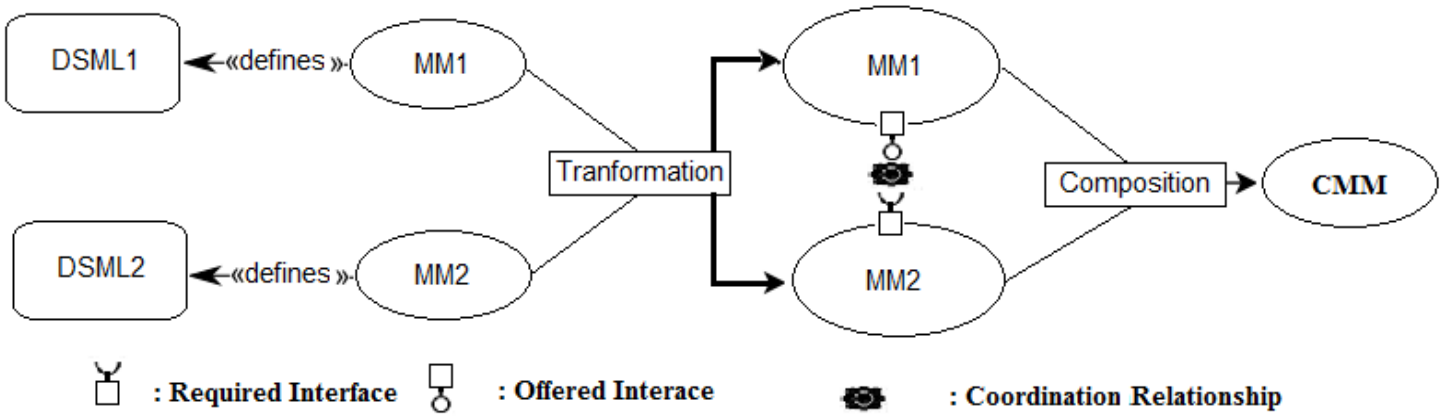


Figure 14: Step1: Coordination Meta-model (CMM) Elaboration

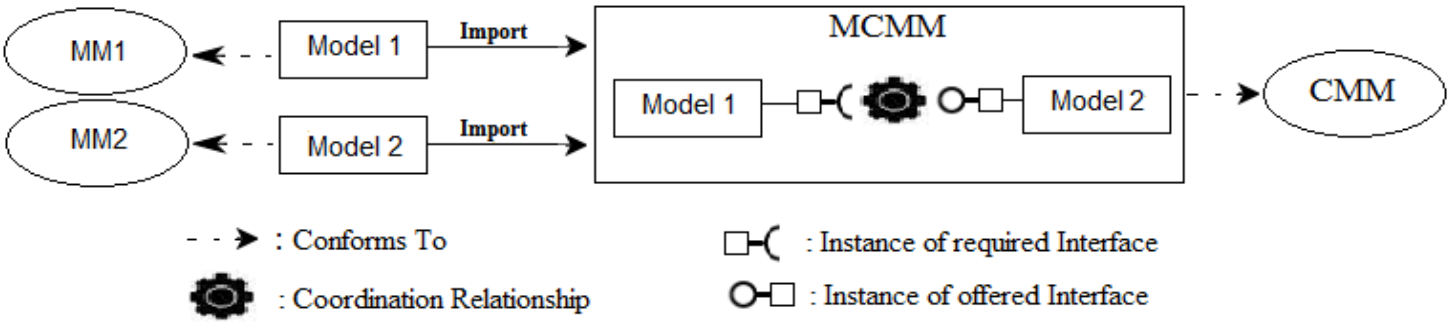


Figure 15: Step2: Models Coordination

5.1. ITS DSML Required Interface

The main feature of ITS DSML is the modeling of an Indoor Transport Service.

However, ITS team needs to get a real time activity of ITSRobot as well as position and battery thresholds. They also need to get location of Item and Room concepts. As said earlier, needs of a DSML are translated to required interface, for that we propose to define a required interface for ITS DSML. This interface is represented by the abstract meta-class named "ConnectedElement" see Figure 16.

According to Bridge Design Pattern, this meta-class is considered as super class of ITSRobot, Item and Room concepts. Moreover, this meta-class must use an element that provides its needs. The used element must be defined by another DSML offered interface. In the current example, the IoT DSML represents the other language.

5.2. IoT DSML Offered Interface

The IoT DSML offers connecting ability to objects by relating them to sensors able to detect events upon them. Subsequently and according to Bridge Pattern design, we propose to define the interface "IObject" as offered interface see Figure 17. This interface is implemented by Object concept. Actually, referencing "IObject" enables connection to sensors.

5.3. CMM Elaboration

The coordination meta-model (CMM) of ITS and IoT DSMLs assures the composition of the two languages (Figure 18).

It has as root "ITS_IOT" element that contains involved languages root elements "ITS" and "IOT" for instance.

According to proposed approach, the coordination between the two languages must be done by defining coordination relationships between both required interface of ITS and offered interface of IoT, described in earlier paragraph. Subsequently, 'ConnectableElement' the super class of both 'ITSRobot', 'Item' and 'Room' must reference the interface 'IObject' of IoT DSML. This later is implemented by 'Object' meta-class belonging to IoT DSML.

While coordination elaborated between ITS and IoT DSMLs is qualified as external [25], we notice that relationship used between those languages interfaces is a composition relationship and more specifically a referencing relationship [25]: "ConnectedElement" references "Iobject".

While coordination elaborated between ITS and IoT DSMLs is qualified as external [25], we notice that relationship used between those languages interfaces is a composition relationship

and more specifically a referencing relationship [25]: "ConnectedElement" references "Iobject".

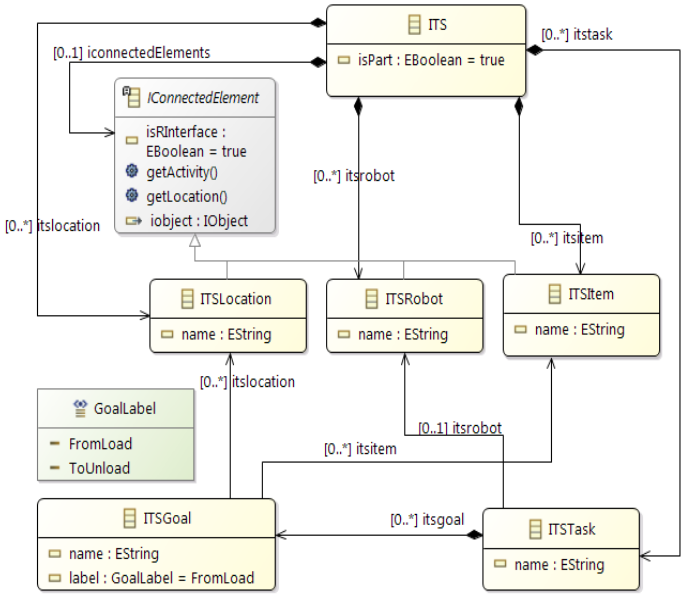


Figure 16: ITS Required Interface

The last step of the coordination process is the coordination between respective Models.

Figure 19 illustrates an ITS Model of a hospital robot transporting medical items from storage to patient rooms. This model has been linked to two IoT models.

The 'Hospital-Robot' has the ability to reference an Object due to the previous elaborated coordination. Hence, we associate it to a device Object having two sensors: Speed sensor to get activity and state of the 'Hospital-Robot and the Battery sensor to monitor battery thresholds of the 'Hospital-Robot.

On the other side, The 'PatientRoom' and 'Storage" have also the ability to be related to an Object element. Thereby, it has been associated to 'Building' object.

These associations provide connection to a Location sensor that gives information about Patient rooms and storage localization.

The coordination done at language level is also valuable for all conformed models. Thus, the same CMM could be used to coordinate models belonging to coordinated DSMLs. Consequently, to supervise elements of the second model illustrated in Figure 6 and Figure 7, we use the same CMM. Thereby, in Figure 20 'FoodDeliveryRobot' and 'CustomerAddress' have both the possibility to reference an 'Object' element of IoT DSML.

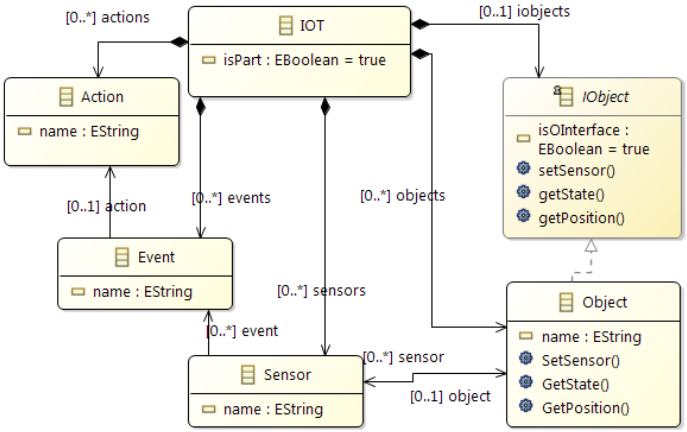


Figure 17: IoT Offered Interface

6. Related Works

Coordination between heterogeneous DSMLs is an active research topic. Although, many research papers discussed this subject, their objectives and motivations are different. Most of them emphasize reutilization and development of DSMLs.

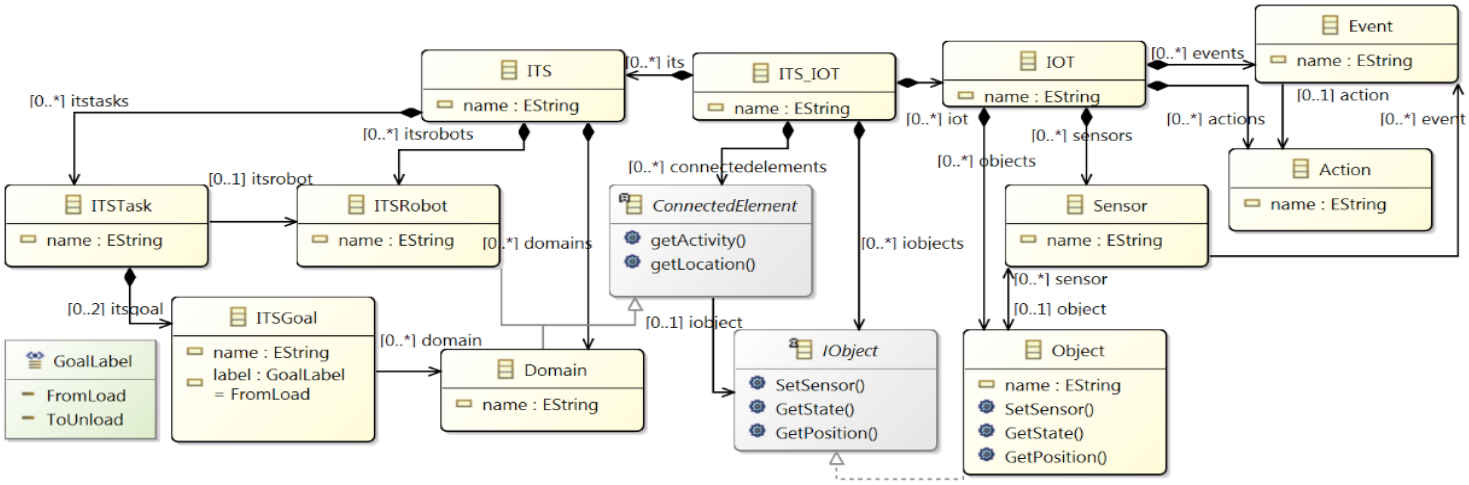


Figure 18: IoT Offered Interface

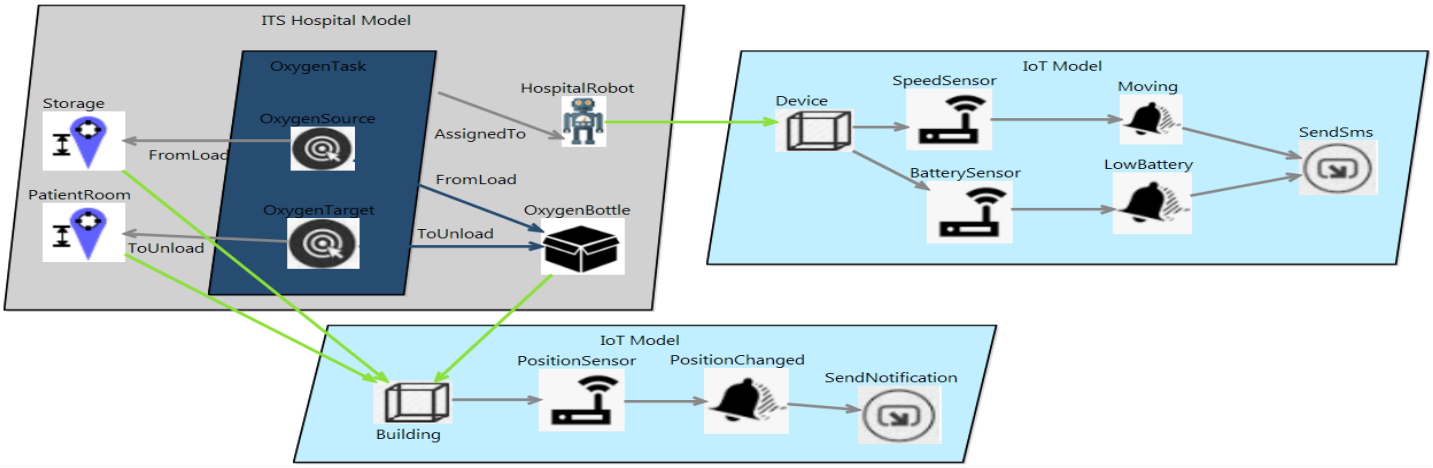


Figure 19: ITS and IoT Models Coordination

This coordination aims to help teams to cooperate with each other in different phases of systems implementation: specification, development, test, maintenance and evolution.

Modularization and interface-based techniques to compose and integrate DSMLs has been proposed by recent papers. A.Klepee [10] discussed the idea of language interfaces to combine languages and like this paper the author gave also a way to define a language interface using attributes and methods. However, our proposition is more explicit. Either, in [8] authors introduced an interface-based approach to define modular meta-models. The approach advises the use of a black-box meta-model definition based on interfaces and corresponding composition operators. This is done using an abstract meta-modeling language enabling the applicability of the approach for various meta-modeling languages. In the same line, our work proposes a composition technique based on offered and required interfaces.

In accordance with this work, Paper [11] advocates the definition of provided and required interfaces for languages as well as their syntactical and semantical composition operators to support DSLs modularity. However, author doesn't talk about how interfaces must be defined. In contrast, in [35] author proposes "Model Types" as example of DSLs interfaces defined on the abstract syntax of languages. Here, models are linked to Model Types using a typing relationship, while Model Types are related to each other with a subtyping relationship.

In [9], author proposes to coordinate heterogeneous behavioral models by specifying coordination patterns between languages. This is done using language behavioral interfaces as well as correspondence and coordination rules. Language behavioral interface in this approach is a set of Domain Specific Events (DSEs) defined in the context of language syntax meta-classes. DSEs are used as coordination points allowing both observation and control of models execution. Actually, coordination is enrolled using BCool the coordination language. That language enables the definition of operators containing matching and coordination rules over DSEs. This proposition is similar to ours as language interface is defined on top of language's abstract syntax even if author didn't give a precise design for interface specification.

In [36] Marco Di Natal et al. use a mapping language to integrate a functional language and a platform language. For that, a set of connectors are used to specify the mapping between the two languages. Once the mapping model is elaborated, communication code between functional model and platform model is generated. We can say that there are some similarities between this proposition and ours. Connectors and mapping model play respectively the same role of interface and coordination model in our proposition.

Finally, in [37] authors define composition interface for model fragments as a set of ports representing reference and variation points of a language. Reference points are root nodes for model fragments, where variation points are nodes that may be replaced during composition. Ports have unique names and map to one or more elements of the fragment. This work is interesting as we ca consider model fragments as separate models that authors coordinate using ports.

7. Conclusion and Future Works

The use of divers DSMLs for modeling a same complex system becomes a common practice. Indeed, systems are more and more complex inciting experts to use their own specific modeling languages.

Albeit, while many language workbenches exist and allow an easy creation of DSMLs, the coordination and integration between them is not proposed or poorly provided.

Subsequently, many recent researches discussed this issue. In the same line, this paper introduces an approach based on both composition and interface concepts to coordinate heterogeneous DSMLs.

Actually, a detailed process composed by two main steps has been proposed. The first step consists in composing involved DSMLs and adds an interface layer to them. This layer is defined according to Bridge Pattern Design.

The proposed approach has been applied to a connected Indoor Transport Service system. The former involves two heterogeneous DSMLs.

For our future works, we plan to explore more relevant cases involving various coordination relationships. Hence, further case studies will help us to improve our coordination process. We also plan to provide a wizard enabling the use of proposed approach and assisting integrators to achieve coordination.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Acknowledge your institute/ funder.

References

[1] M. Fowler, R. Parsons, Domain Specific Languages, Addison-Wesley Professional, 2010.
[2] C. Hardebolle, " Composition de modèles pour la modélisation multi-paradigme du comportement des systems," Ph.D Thesis, Université Paris-Sud XI Orsay, 2008.
[3] D. Simon-Zayas, "A framework for the management of heterogeneous models in Systems Engineering,", Ph.D Thesis, Ecole Nationale Supérieure de Mécanique et d'Aérotechnique, 2012.
[4] M. El hamlaoui, "Mise en correspondence et gestion de la cohérence de modèles hétérogènes évolutifs," Ph.D Thesis, Université Toulouse-Jean-jaurès, 2015.
[5] M. Chechik, S.Nejati, M.Sabetzadeh, "A Relationship-Based Approch to Model Integration" Innovations in Systems and Software Engineering, 8(1), 3-18, 2012. https://doi.org/10.1007/s11334-011-0155-2
[6] B. Combemale, J. Deantoni, B. Baudry, J. Jézéquel and J. Gray, "Globalizing Modelling Languages" Computer 47.6:68-71, 2014. https://doi.org/ 10.1109/MC.2014.147
[7] J. Deantoni, C. Brun, B. Caillaud, R.B. France, G. Karsai, O. Nierstrasz and E. Syriani, "Domain globalization: using languages to support technical and social coordination", Globalizing Domain-Specific Languages (pp. 70-87). Springer, Cham, 2015. https://doi.org/10.1007/978-3-319-26172-0_5
[8] S. Živković and D. Karagiannis, "Towards metamodelling-in-the-large: Interface-Based composition for modular metamodel development" in Enterprise, Business-Process and Information Systems Modeling. Springer, Cham, 2015. https://doi.org/10.1007/978-3-319-19237-6_26
[9] M.E.V Larsen, J. Deantoni, B. Combemale and F. Mallet, "A Behavioral Coordination Operator Language (BCOol)" in Model Driven Engineering Languages and Systems (MODELS), ACM/IEEE 18th International Conference, 2015. https://doi.org/10.1109/MODELS.2015.7338249
[10] A. Klepee, Software language engineering, Addison-Wesley Professional, 2008.
[11] T. Degueule, "Towards Language Intefaces for DSLs Integration", 2015. https://hal.inria.fr/hal-01138017
[12] D. Strüber, G. Taentzer, S. Jurack & T. Schäfer. "Towards a distributed modeling process based on composite models" in International Conference on Fundamental Approaches to Software Engineering. Springer, Berlin, Heidelberg, 2013. https://doi.org/10.1007/978-3-642-37057-1_2
[13] D. Strüber, S. Jurack,, T. Schäfer, S. Schulz & G. Taentzer. "Managing Model and Meta-Model Components with Export and Import Interfaces." in Big- MDE Workshop on Scalability in Model Driven Engineering. p. 31-36, 2016
[14] ITU-T Recommendation X.208, "Specification of Abstract Syntax Notation One (ASN.1)", 1988.
[15] ITU-T Recommendation X.209, "Specification of Basic Encoding Rules for Abstract Syntax Notation One (ASN.1)", 1988.
[16] M.Bjorklund, "YANG a data modeling language for the network configuration protocol (NETCONF)" (No. RFC 6020), 2010.
[17] M. Bjorklund, "The YANG 1.1 Data Modeling Language (No. RFC 7950)", 2016.
[18] ISO/IEC, "Standard EBNF Syntaxt Specification". Ebnf: Iso/iec 14977, 1996 (e). URL http://www. cl. cam. ac. uk/mgk25/iso-14977. pdf, 1996, vol. 70.
[19] J. Paakki. Attribute grammar paradigms—a high-level methodology in language implementation. ACM Computing Surveys (CSUR), vol. 27, no 2, p. 196-255, 1995.
[20] G. Rozenberg. "Handbook of Graph Grammars and Comp", World scientific, 1997.
[21] L. Fuentes-Fernández & A. Vallecillo-Moreno, "An introduction to UML profiles". UML and Model Engineering, vol. 2, p. 6-13, 2004.
[22] T. Kühne, "Matters of (meta-) modeling". Software & Systems Modeling, vol. 5, no 4, p. 369-385, 2006. https://doi.org/10.1007/s10270-006-0017-9
[23] T. Clark, P. Sammut and J. Willans, Applied metamodeling: a foundation for language driven development, Middlesex University Research Repository, 2008.
[24] K. Chen, J. Sztipanovits and S. Neema, "Toward a semantic anchoring infrastructure for domain-specific modeling languages" in the 5th ACM international conference on Embedded software (pp. 35-43). 2005. https://doi.org/10.1145/1086228.1086236
[25] M. Klein, "Combining and relating ontologies: an analysis of problems and solutions". in Workshop on Ontologies and Information Sharing, IJCAI, 2001
[26] J. P. Tolvanen, R. Pohjonen & S.Kelly. " Advanced tooling for domain-specific modeling: MetaEdit+ " in the 7th OOPSLA Workshop on Domain-Specific Modeling, Finland, 2007. ISBN: 978-951-39-2915-2
[27] R. Heim, P.M.S Nazari, J.O. Ringert, B. Rumpe and A. Wortmann, "Modeling robot and world interfaces for reusable tasks" in IEEE Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on (pp. 1793-1798), 2015. https://doi.org/10.1109/IROS.2015.7353610
[28] E. Juliot and J. Benois, Viewpoints creation using obeo designer or how to build eclipse dsm without being an expert developer?, Obeo, Tech. Rep., obeo designer whitepaper, 2010. http://www.obeo.fr
[29] Graphical Editing Framework (GEF), http://www.eclipse.org/gef/.
[30] Eclipse Community Forums: GMF (Graphical Modeling Framework): http://www.eclipse.org/forums/eclipse.modeling.gmf.
[31] K. Hölldobler, A. Roth, B. Rumpe and A. Wortmann, "Advances in modeling Language engineering" in Proc of International Conference on Model and Data Engineering (pp. 3-17),Springer,2017. https://doi.org/10.1007/978-3-319-66854-3_1
[32] T. Deguele, "Composition and interoperability for external domain-specific language engineering", Ph.D Thesis, Rennes 1,2016.
[33] T. Clark, M. Van den Brand, B. Combemale and B. Rumpe, "Conceptual model of the globalization for domain-specific languages" in Globalizing Domain-Specific Languages(pp.7-20). Springer, Cham, 2015. https://doi.org/10.1007/978-3-319-26172-0_2
[34] E. Gamma, R. Helm, R. Johnson, J. Vlissides, Design patterns: Elements of Reusable Object-Oriented Software, Pearson Education India ,1995.
[35] J. Steel, J.M Jézéquel, "On model typing" Software & Systems Modeling 6(4), 401–413, 2007. https://doi.org/10.1007/s10270-006-0036-6
[36] M. Di Natale, F. Chirico, A. Sindico and A. Sangiovanni- Vincentelli, "An MDA approach for the generation of communication adapters integrating SW and FW components from Simulink" in International Conference on Model Driven Engineering Languages and Systems (pp. 353-369), Springer, Cham, 2014. https://doi.org/10.1007/978-3-319-11653-2_22
[37] J. Johannes, R. Samlaus and M. Seifert, "Round-trip support for invasive software composition systems" in July 2009 Springer, Berlin, Heidelberg, International Conference on Software Composition (pp. 90-106), 2009. https://doi.org/10.1007/978-3-642-02655-3_8

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 150-157 (2020)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Non Parallelism and Cayley-Menger Determinant in Submerged Localization

Anisur Rahman*

Department of Computer Science and Engineering, East West University, Dhaka – 1212, Bangladesh

ARTICLE INFO

Article history:
Received: 16 January, 2020
Accepted: 22 April, 2020
Online: 08 May, 2020

Keywords:
Cayley-Menger determinant
Linearization
Mobile beacon
Parallel and non-parallel state
Submerged localization
Underwater wireless sensor network

ABSTRACT

This research paper portraits the technique to determine location of submerged nodes with Cayley-Menger determinant and associated problems with non-parallel states. Cayley-Menger determinant is considered to be the usual means to determine the coordinates of the nodes with a single node where the plane of beacon i.e. beacon's surfing plane and the plane created by the deployed submerged sensors are in parallel state. However, this perfect scenario hardly exists in the submerged world; as a result Cayley-Menger determinant may not be used unless proper measurements are taken. This paper addresses this limitation and analyzed the proposed model to deal with this non-parallel state situation. As precise localization is vital for the validity of the sensed data in Underwater Wireless Sensor Networks (UWSNs); exact distance measurement between nodes is an important and crucial in range based localization methods. Proposed method has addressed uncertainty of nonlinear equations as well as how better immunity could be achieved in multipath fading in propagation. Analyzed simulation and experimental results conclude the accuracy of the proposed model with minimal error, where it has been shown that parallelism of the system is not a factor for using Cayley-Menger determinant. Moreover, a single node has been used in the model to localize submerged nodes where none acquires priori familiarity about its position.

1. Introduction

Underwater wireless sensor networks (UWSNs) are envisioned for exploring the vast underwater world for the profusion of wealth and mystery. Not only the wealth of underwater world, submerged localization necessary for the sustenance of our very own existence and for the marine biome. Lately researchers have shown fervent importance to investigate, explore and analyze the abundance of underwater world; so it has become indispensable to collect accurate environmental data with the help of underwater sensors. Exact localization is not only necessary for underwater applications; it determines the very nature of our own existence at rudimentary and core level as well as helps us to provide a sanctuary for the marine biome. Moreover, autonomous underwater vehicles (AUVs) control and surveillance, monitoring seabed and faults for upcoming natural calamity, searching for lost object, pollutants and nutrients tracking also demand precise localization [1]. Among these applications, some significantly requires accurate localization for meaningful comprehension of gathered data for a practical use [2].

*Anisur Rahman, Dept. of Computer Sci. & Engg., EWU, anis@ewubed.edu

UWSNs may comprise of deployed sensors and surface stations; many a time localization of submerged sensors is done with the help of multiple surfaced nodes. However, localizing using a single beacon is considered to be pragmatic as it requires less provision. Despite numerous underwater applications, the concept and achievement of underwater wireless communication may still seem far-fetched. Mostly, communications is done with acoustic signals as radio signals have very limited propagation in underwater; hence, using acoustic signals for distance measurements in range based localization provides more accuracy that using radio [1,3]. However, multipath fading and synchronization between communicating nodes still remain to be the challenging factors so far.

Duff and Muller solved system of multilateration equations by applying nonlinear square optimization method where positions of the sensors remain unknown [3]. This proposed algorithm is solely based on degree-of-freedom analysis phenomenon, which dictates sufficient numbers of measurements are necessary to generate adequate number of equations for the problem to solve. This proposed method has been validated and showed the improved accuracy by many folds using Kalman filter in [4]. Incorporation

of Kalman filter increases the complexity of the method with respect to degree-of-freedom complexity of system of equations in [5]. It also delineates that the method does not guarantee a unique solution for system of non-linear equations, i.e. trilateration. Moreover, the method requires a specific initial configuration of the nodes which was justified by rigidity theory. 3D positioning system in [6] requires four separate positions to determine the coordinates of the beacon.

This paper analyzed the method proposed in [7] to determine the coordinates of underwater sensors with a single beacon where Cayley-Menger determinant is used for non- parallel situation and validates the method with simulation and experimentation. Recently, localization of submerged sensors for non-parallel states has been proposed in [8]. It also showed that the coordinates could be determined in adhoc basis without any pre-installed infrastructure. Moreover, it has established that the accuracy of the coordinates depends on the distance measurements between surfaced node and underwater deployed sensors, not the state of the planes whether parallel or non-parallel.

The arrangement of the remaining paper is as following where Section 2 focuses on mathematical model of coordinates determination both for parallel and non-parallel state situations; it also talks about linear transformation of the origin to find the coordinates with respect to sensor and beacon separately. Section 3 states simulation to validate the proposed mathematical model, experimental results also shows the accuracy of the model and at last analysis. Section 4 states conclusion with future works.

2. Coordinates Determination

The proposed algorithm in [9] was to determine the coordinates of the submerged sensors with a single beacon; the problem domain was considered to be in parallel state which is only possible in perfect world. However, in [7] Cayley-Menger determinant has been used for non-parallel state situation. Following sections iterates the proposed algorithm for parallel and non-parallel states.

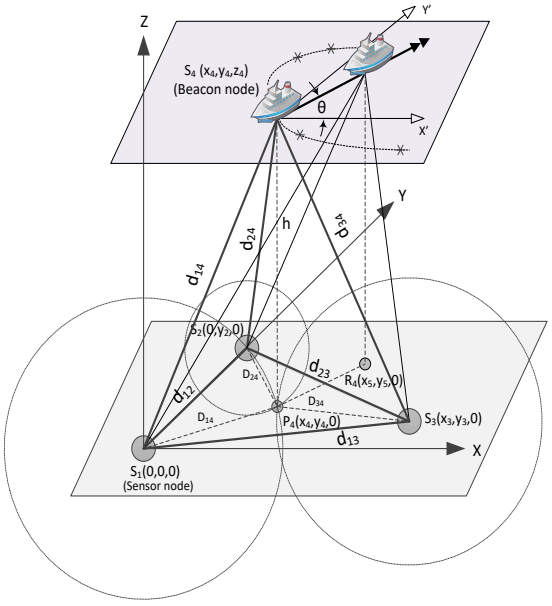


Figure 1: Coordinates Determination

2.1. Coordinates Determination (parallel state)

The problem domain consists of a single beacon $S_j, j = 4,5....9$ which is on the water surface and three underwater sensors $S_i, i = 1,2,3$ whose coordinates need to be determined. For simplicity, one of the sensors is considered to the Cartesian origin $(0,0,0)$ and one other placed on another axis leaving the third one on the same quadrant. The distance between the beacon $S_j, j = 4,5....9$ and the deployed sensors $S_i, i = 1,2,3$ have been computed according to the process depicted in [10], these measured values are $d_{14}, d_{24}, d_{34}....$ and unknown inter nodes distances are d_{12}, d_{13}, d_{23}. From these values, volume of the tetrahedron V_t which is created by the beacon and the deployed sensors depicted in Figure 1can be written with Cayley-Menger determinant as in (1).

$$288\,V_t^2 = \begin{vmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & d_{12}^2 & d_{13}^2 & d_{14}^2 \\ 1 & d_{12}^2 & 0 & d_{23}^2 & d_{24}^2 \\ 1 & d_{13}^2 & d_{23}^2 & 0 & d_{34}^2 \\ 1 & d_{14}^2 & d_{24}^2 & d_{34}^2 & 0 \end{vmatrix} \tag{1}$$

By expanding (1), we obtain:

$$d_{34}^2 d_{23}^2 - d_{34}^2 d_{12}^2 + d_{34}^2 d_{13}^2 - \frac{d_{14}^2 d_{23}^4}{d_{12}^2} + d_{23}^2 d_{14}^2 + \frac{d_{13}^2 d_{14}^2 d_{23}^2}{d_{12}^2} - \frac{d_{24}^2 d_{13}^4}{d_{12}^2} + \frac{d_{13}^2 d_{23}^2 d_{24}^2}{d_{12}^2} + d_{13}^2 d_{24}^2 - d_{13}^2 d_{23}^2 - 144\frac{V_t^2}{d_{12}^2} + \frac{d_{14}^2 d_{23}^2 d_{24}^2}{d_{12}^2} + \frac{d_{14}^2 d_{23}^2 d_{34}^2}{d_{12}^2} - \frac{d_{23}^2 d_{24}^2 d_{34}^2}{d_{12}^2} - \frac{d_{14}^4 d_{23}^2}{d_{12}^2} + \frac{d_{13}^2 d_{24}^2 d_{34}^2}{d_{12}^2} - \frac{d_{13}^2 d_{14}^2 d_{34}^2}{d_{12}^2} + \frac{d_{13}^2 d_{14}^2 d_{24}^2}{d_{12}^2} - \frac{d_{13}^2 d_{24}^4}{d_{12}^2} - d_{34}^4 + d_{24}^2 d_{34}^2 + d_{14}^2 d_{34}^2 - d_{14}^2 d_{24}^2 = 0$$

By isolating and grouping known and unknown variables:

$$d_{34}^2\left(d_{12}^2 - d_{23}^2 - d_{13}^2\right) + d_{14}^2\left(\frac{d_{23}^4}{d_{12}^2} - d_{23}^2 - \frac{d_{13}^2 d_{23}^2}{d_{12}^2}\right) + d_{24}^2\left(\frac{d_{13}^4}{d_{12}^2} - \frac{d_{13}^2 d_{23}^2}{d_{12}^2} - d_{13}^2\right) - \left(d_{14}^2 d_{24}^2 + d_{14}^2 d_{34}^2 - d_{24}^2 d_{34}^2 - d_{14}^4\right)\frac{d_{23}^2}{d_{12}^2} - \left(d_{34}^2 d_{24}^2 - d_{14}^2 d_{34}^2 + d_{14}^2 d_{24}^2 - d_{24}^4\right)\frac{d_{13}^2}{d_{12}^2} + \left(144\frac{V_t^2}{d_{12}^2} + d_{13}^2 d_{23}^2\right) = \left(d_{24}^2 d_{34}^2 - d_{34}^4 + d_{14}^2 d_{34}^2 - d_{14}^2 d_{24}^2\right)$$

Here, $\left(\frac{d_{13}^4}{d_{12}^2} - \frac{d_{13}^2 d_{23}^2}{d_{12}^2} - d_{13}^2\right)$, $\left(\frac{d_{23}^4}{d_{12}^2} - d_{23}^2 - \frac{d_{13}^2 d_{23}^2}{d_{12}^2}\right)$, $\frac{d_{23}^2}{d_{12}^2}$, $\left(d_{12}^2 - d_{23}^2 - d_{13}^2\right)$, $\frac{d_{13}^2}{d_{12}^2}$, and $\left(144\frac{V_t^2}{d_{12}^2} + d_{13}^2 d_{23}^2\right)$ are unknowns.

Here, we rewrite the equation in the following form:

$$d_{14}^2 X_1 + d_{24}^2 X_2 + d_{34}^2 X_3 - \left(d_{14}^2 - d_{34}^2\right)\left(d_{24}^2 - d_{14}^2\right)X_4 - \left(d_{24}^2 - d_{14}^2\right)\left(d_{34}^2 - d_{24}^2\right)X_5 + X_6 = \left(d_{24}^2 - d_{34}^2\right)\left(d_{34}^2 - d_{14}^2\right) \tag{2}$$

The above equation has six unknowns as depicted by X_1, X_2, X_3, X_4, X_5 and X_6; and the equation resembles with the linear form of

$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b_1.$

S_4, so that in real life the movement would be less between distance measurements. From these six equations reference to the variables have been omitted to get an array of all coefficients, which is represented as augmented matrix. The matrix consists of linear equations and resembles $AX = b$ formula. So, the system of linear equations is articulated as following:

$$A = \begin{bmatrix} d_{14}^2 & d_{24}^2 & d_{34}^2 & -\left(d_{14}^2 - d_{34}^2\right)\left(d_{24}^2 - d_{14}^2\right) & -\left(d_{24}^2 - d_{14}^2\right)\left(d_{34}^2 - d_{24}^2\right) & 1 \\ d_{15}^2 & d_{25}^2 & d_{35}^2 & -\left(d_{15}^2 - d_{35}^2\right)\left(d_{25}^2 - d_{15}^2\right) & -\left(d_{25}^2 - d_{15}^2\right)\left(d_{35}^2 - d_{25}^2\right) & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ d_{19}^2 & d_{29}^2 & d_{39}^2 & -\left(d_{19}^2 - d_{39}^2\right)\left(d_{29}^2 - d_{19}^2\right) & -\left(d_{29}^2 - d_{19}^2\right)\left(d_{39}^2 - d_{29}^2\right) & 1 \end{bmatrix}$$

$$X = \begin{bmatrix} \left(\frac{d_{23}^4}{d_{12}^2} - d_{23}^2 - \frac{d_{13}^2 d_{23}^2}{d_{12}^2}\right) \\ \left(\frac{d_{13}^4}{d_{12}^2} - \frac{d_{13}^2 d_{23}^2}{d_{12}^2} - d_{13}^2\right) \\ \left(d_{12}^2 - d_{23}^2 - d_{13}^2\right) \\ \frac{d_{23}^2}{d_{12}^2} \\ \frac{d_{13}^2}{d_{12}^2} \\ \left(144\frac{V_t^2}{d_{12}^2} + d_{13}^2 d_{23}^2\right) \end{bmatrix} \qquad b = \begin{bmatrix} \left(d_{24}^2 - d_{34}^2\right)\left(d_{34}^2 - d_{14}^2\right) \\ \left(d_{25}^2 - d_{35}^2\right)\left(d_{35}^2 - d_{15}^2\right) \\ \vdots \\ \left(d_{29}^2 - d_{39}^2\right)\left(d_{29}^2 - d_{19}^2\right) \end{bmatrix}$$

From the above representation, after finding X_1, X_2, X_3, X_4, X_5 and X_6 we calculate d_{12}, d_{13} and d_{23} as follows:

$$d_{12}^2 = \frac{X_3}{(1 - X_4 - X_5)},\ d_{13}^2 = \frac{X_3 X_5}{(1 - X_4 - X_5)},\ d_{23}^2 = \frac{X_3 X_4}{(1 - X_4 - X_5)}$$

Coordinates of deployed submerged sensors $S_1(0,0,0)$ is considered to be the origin of the Cartesian system, $S_2(0, y_2, 0)$ is on one of the *y-axis* and $S_3(x_3, y_3, 0)$ is considered on some position not on the axis as depicted in Figure 1. From the above equations, the inter node distances can be calculated as follows:

$$d_{12}^2 = y_2^2,\ d_{13}^2 = x_3^2 + y_3^2\ ,\ d_{23}^2 = x_3^2 + (y_3 - y_2)^2$$

As a result, the unknown variables y_2, y_3 and x_3 can be derived with respect to inter node distances:

$$y_2 = d_{12},\ y_3 = \frac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d_{12}},\ x_3 = \sqrt{\left(d_{13}^2 - \left(\frac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d_{12}}\right)^2\right)}$$

Here, d_{12}, d_{13} and d_{23} are computed distances between deployed submerged sensors and Table 1 shows the coordinates of the sensors as follows:

Table 1: Calculated Coordinates of the Submerged Sensors

Sensors	Coordinates
S_1	$(0,0,0)$
S_2	$(0, d_{12}, 0)$
S_3	$\left(\sqrt{\left(d_{13}^2 - \left(\frac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d_{12}}\right)^2\right)}, \frac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d_{12}}, 0\right)$

2.2. Coordinates with respect to Beacon (parallel state)

So far the previous section has dealt with determination of coordination when beacon's surfing plane and plane of deployed submerged sensors are in parallel. Once the transformation of the origin takes place and taken to the beacon's position, the localization process would be following.

In this method, a pressure sensor is used to measure the depth *h* of Figure 2, which is elaborated in [11]. Once vertical distance (*h*) between the planes is known, the projected coordinates of the beacon $S_4(x_4, y_4, z_4)$ can be found on the *XY* plane and can be denoted as $P_4(x_4, y_4, 0)$. Trilateration can be used to find x_4 and y_4 of the projected coordinates considering D_{14}, D_{24} and D_{34} to be the distances between S_1, S_2, S_3 and P_4 respectively. Once the projected coordinates and the distances are determined, following relations can be devised.

$$D_{14}^2 = x_4^2 + y_4^2 \quad (3)$$

$$D_{24}^2 = x_4^2 + (y_4 - y_2)^2 \quad (4)$$

$$D_{34}^2 = (x_4 - x_3)^2 + (y_4 - y_3)^2 \quad (5)$$

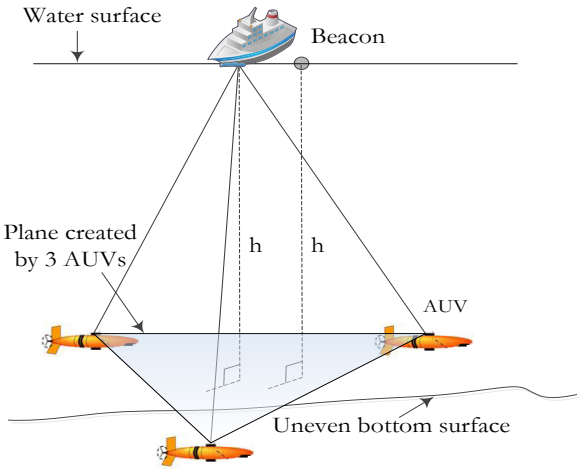


Figure 2: Parallel Plane State Scenario

Equations (3), (4) and (5) help to determine the coordinates of the projected beacon $P_4(x_4, y_4, 0)$, where

$$x_4 = \frac{1}{2d_{12}}\sqrt{\left(4d_{12}^2 D_{14}^2 - (D_{14}^2 - D_{24}^2 + d_{12}^2\right)^2},$$

$$y_4 = \frac{1}{2d_{12}}\left(D_{14}^2 - D_{24}^2 + d_{12}^2\right)$$

Here, d_{14} , d_{24} and d_{34} are the hypotenuse of the $\Delta S_1P_4S_4$, $\Delta S_2P_4S_4$ and $\Delta S_3P_4S_4$ respectively, so it is possible to obtain D_{14}, D_{24} and D_{34} using simple Pythagorean Theorem. As a result, the coordinate of the beacon node that is on the surface, $S_4(x_4, y_4, z_4)$ would be $S_4(x_4, y_4, h)$; here both x_4 and y_4 have been calculated accordingly and h is known from the installed pressure sensor. So, the coordinates of S_4 would be as follows:

$$\therefore S_4(x_4,y_4,h) = S_4\begin{pmatrix} \frac{1}{2d_{12}}\sqrt{\left(4d_{12}^2D_{14}^2 - (D_{14}^2 - D_{24}^2 + d_{12}^2\right)^2}, \\ \frac{1}{2d_{12}}\left(D_{14}^2 - D_{24}^2 + d_{12}^2\right), h \end{pmatrix}$$

Once the origin $(0,0,0)$ of the Cartesian system is transferred on the beacon $S_4(x_4, y_4, h)$, then the coordinates of the deployed sensors with respect to beacon are calculated by linear transformation as depicted in [10].

Table 2: Coordinates of the Submerged Sensors with respect to Beacon for Parallel State

Sensors	Coordinates
S_4	$(0,0,0)$
S_1	$\left(\frac{\sqrt{4d_{12}^2D_{14}^2 - \left(D_{14}^2 - D_{24}^2 + d_{12}^2\right)^2}}{2d_{12}}, \ -\frac{1}{2d_{12}}\left(D_{14}^2 - D_{24}^2 + d_{12}^2\right), \ -h\right)$
S_2	$\left(\frac{\sqrt{4d_{12}^2D_{14}^2 - \left(D_{14}^2 - D_{24}^2 + d_{12}^2\right)^2}}{2d_{12}}, \ \frac{1}{2d_{12}}\left(d_{12}^2 - D_{14}^2 + D_{24}^2\right), \ -h\right)$
S_3	$\left(\left(\sqrt{d_{13}^2 - \left(\frac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d_{12}}\right)^2}\right) - \frac{\sqrt{4d_{12}^2D_{14}^2 - \left(D_{14}^2 - D_{24}^2 + d_{12}^2\right)^2}}{2d_{12}}\right), \ \frac{1}{2d_{12}}\left(d_{13}^2 - d_{23}^2 - D_{14}^2 + D_{24}^2\right), \ -h$

2.3. Coordinates with respect to Beacon (non-parallel state)

Having the plane where the beacon moves around i.e. water surface and the plane that is created by the three deployed sensors in parallel state can be found in very limited situations. Previous coordinates determination method could be adjusted to meet all the situations both parallel and non-parallel. As a result, a derivative of parallel state method has been illustrated in [8]; however this paper also shows the relationship between Cayley-Menger and state of the planes. With a little adjustment as depicted in this section where Figure 3 illustrates the scenario when both planes are in non-parallel state situation.

$$Plane\ A: \prod\nolimits_{S_1,S_2,S_3} : (a_1x + b_1y + c_1z + \delta_1 = 0)$$

Here, plane A is created by the submerged deployed sensors S_1, S_2, and S_3 as three dots (sensors) can create a plane. Whereas,

$$Plane\ B: \prod\nolimits_{Beacon} : (a_2x + b_2y + c_2z + \delta_2 = 0)$$

Here, plane B is where the beacon moves around while taking six distance measurements i.e. water surface, are in non-parallel state.

Figure 3 is the most likely the situation found in the world, except in some special cases like, water tank, swimming pools or where the deployed sensors can maintain a predefined height like AUVs or UUVs as in Figure 2. When we consider the water surface as the reference plane then any deployed node having a different height than others would create a non-parallel state situation. The system of equations and the linearization procedure devised in previous section are meant to be used in parallel state; as volume of six different tetrahedrons created by deployed submerged sensors and the surfaced beacon are always equal. Once apex of six tetrahedrons on the same parallel plane having the base same i.e. the height of the tetrahedrons are same having the base fixed, the volume would be same. This fundamental theory creates the scope to use the devised equations in previous section. Vertical distances will vary depending on the dihedral angle α accordingly to the (6).

$$\cos\alpha = \frac{\hat{n}_1.\hat{n}_2}{\left|\hat{n}_1\right|\left|\hat{n}_2\right|} = \frac{a_1a_2 + b_1b_2 + c_1c_2}{\sqrt{a_1^2 + b_1^2 + c_1^2}\sqrt{a_2^2 + b_2^2 + c_2^2}} \qquad (6)$$

Here, α dihedral angle; $\hat{n}_1 = (a_1, b_1, c_1)$ and $\hat{n}_2 = (a_2, b_2, c_2)$ are normal vectors to plane A and plane B respectively.

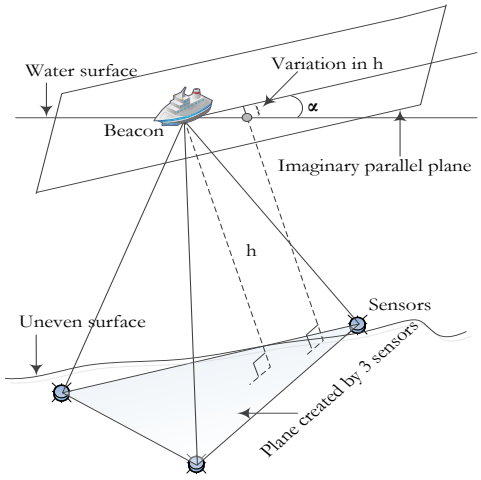


Figure 3: Parallel Planes Effect

It is worth mentioning that all the deployed submerged nodes have several sensors for communication purposes; among them pressure is one of them. Nodes are supposed to read water pressure and communicate to beacon node via acoustic signals where the pressure reading will be converted into depth following method devised in [12]. Once all three sensors' depth is known, following method would be applied to determine the coordinates. Figure 4 illustrates a solvable configuration.

At this point of the procedure all three deployed submerged sensors' depth would be known from the built-in pressure sensors. Let these depth be h_1 h_2 h_3 S_1 , S_2 and S_3

$h_3 > h_2 > h_1$, S_3 is at the lowest point whereas S_1 is at the highest point among all the three sensors with depth h_3 and h_1 respectively. In this situation, it is conspicuous that plane Π_{beacon} where beacon surfs i.e. water surface and the plane $\Pi_{s_1s_2s_3}$ created by three submerged sensors i.e. S_1, S_2 and S_3 would be in non-parallel state.

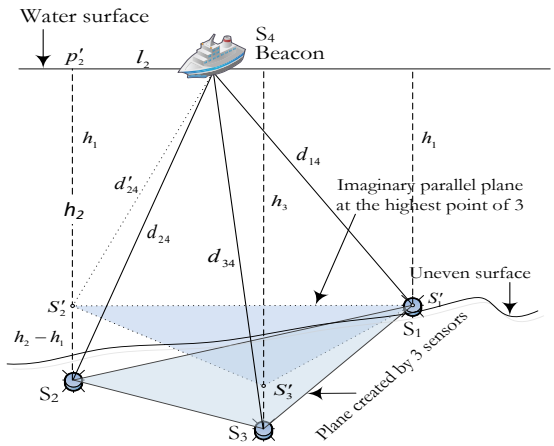


Figure 4: Parallel State in Non-parallel Situation

Let the shortest depth or highest point be S_1', where $S_1' = S_1$, considering S_2' and S_3' are be two points where exactly right above S_2 and S_3 having equal depth with S_1', here the coordinates of S_2' and S_3' would be same as S_2 and S_3 except z component as these points are right above the nodes. As a result, the plane $\Pi_{s_1s_2s_3}$ would be in parallel state with the plane $\Pi_{s_1's_2's_3'}$.

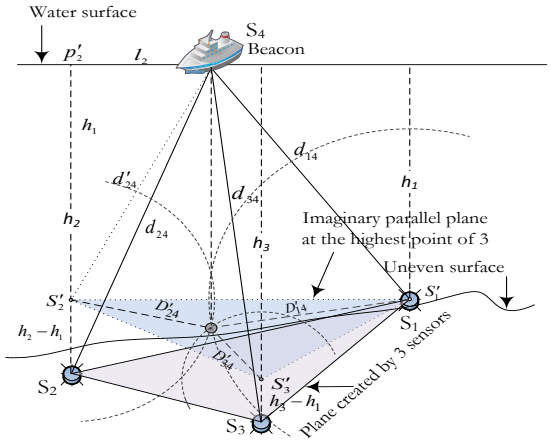


Figure 5: Trilateration and Coordinates Computation

In $\Delta S_4 S_2 p_2'$, where p_2' is the projection point right above S_2, and d_{24} is the measured distance what we get with acoustic signal as delineated in [13]. The distance l_2 can be calculated from $l_2 = \sqrt{d_{24}^2 - h_2^2}$. Once l_2 is known, the distance d_{24}' may be calculated from the $\Delta S_4 S_2' p_2'$ according to (9).

$$d_{24}' = \sqrt{l_2^2 + h_1^2} \tag{7}$$

The aforesaid mentioned technique will be followed to calculate distances between the beacon S_4 and the point S_3'. Once both distances between beacon S_4 to points S_2' and S_3' are determined, coordinates determination procedure will begin for all imaginary points as well as for all the deployed submerged sensors.

Table 2 and 3 show the calculated coordinates for parallel and non-parallel states respectively. So, the devised procedure - Cayley-Menger determinant with a single beacon can be used for both parallel and non-parallel state situations with a little adjustment.

Table 3: Coordinates of the Submerged Sensors with respect to Beacon for Non-parallel State

	Coordinates
S_4	(0,0,0)
S_1	$\left(\frac{\sqrt{4d_{12}'^2 D_{14}'^2 - \left(D_{14}'^2 - D_{24}'^2 + d_{12}'^2\right)^2}}{2d_{12}'}, \; -\frac{1}{2d_{12}'}\left(D_{14}'^2 - D_{24}'^2 + d_{12}'^2\right), \; -h_1 \right)$
S_2	$\left(\frac{\sqrt{4d_{12}'^2 D_{14}'^2 - \left(D_{14}'^2 - D_{24}'^2 + d_{12}'^2\right)^2}}{2d_{12}'}, \; \frac{1}{2d_{12}'}\left(d_{12}'^2 - D_{14}'^2 + D_{24}'^2\right), \; -h_2 \right)$
S_3	$\left(\left(\sqrt{d_{13}'^2 - \left(\frac{d_{12}'^2 + d_{13}'^2 - d_{23}'^2}{2d_{12}'} \right)^2} - \frac{\sqrt{4d_{12}'^2 D_{14}'^2 - \left(D_{14}'^2 - D_{24}'^2 + d_{12}'^2\right)^2}}{2d_{12}'} \right), \; \frac{1}{2d_{12}'}\left(d_{13}'^2 - d_{23}'^2 - D_{14}'^2 + D_{24}'^2\right), \; -h_3 \right)$

3. Simulation and Experimental Results

To validate and analyze the proposed model, a simulation has been conducted in Matlab for an aforesaid problem domain; later a stringent experiment has been performed with ultrasonic sensors in terrestrial environment to reinforce the proposed method.

3.1. Simulation Results

The problem domain imitates a scenario for a 200m water column where a single beacon has been used. The method has been devised in [9] elaborately where a group of three sensors are place arbitrarily and a single beacon is placed to imitate the surfaced sensor. The plane where the sensors have been deployed considered being *XY* plane and the beacon's surfing area has been considered to be parallel to *XY* plane. First sensor's coordinate is (0,0,0), which is considered to be the origin of the Cartesian system; whereas the second sensor's is placed on (0,75,0) being on the y-axis. Lastly, the third sensor was placed on (80,40,0) where z component is zero like others. Other environmental variables for the water column are - 30°C at the surface and 15°C at the bottom where sensors are deployed having a salinity variation of 0.5ppt between surface and *XY* plane. Gaussian noise with (μ=0, σ=1) has been added to bottom temperature as well as with flight time of acoustic signals from beacon to submerged nodes.

As the procedure demands, the surfaced sensors has been shifted and measured from six different places; all these six positions were in close proximity to mitigate error incorporation.

However, mobility of the deployed sensors were ignored in the proposed model as this is out of the scope of this paper. Possible errors in determined coordinate of S_2 and S_3 are shown in Figure 6 and 7 for several iterations. As sensor S_1 has been kept at the origin of Cartesian coordinate system; hence generating no error. Having positional distance error of 0.62m with standard deviation of 0.478 for sensor S_2 and 0.75m with standard deviation of 0.603 for sensor S_3 also validate the model.

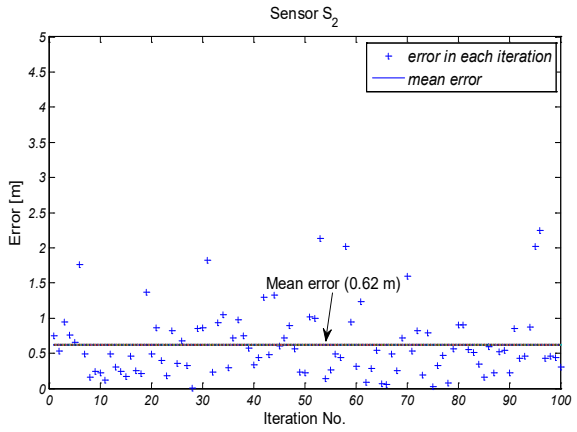


Figure 6: Positional Error from Original Position (0,75,0) for Sensor S_2 (without Gaussian noise)

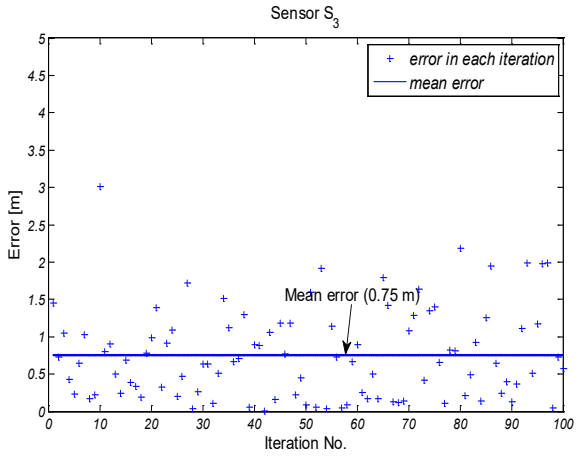


Figure 7: Positional Error from Original Position (80,40,0) for Sensor S_3 (without Gaussian noise)

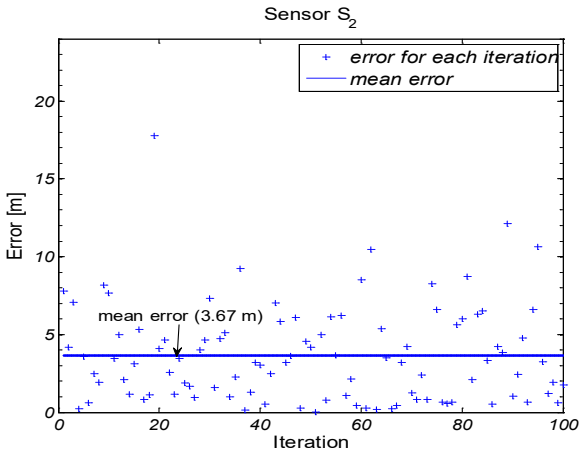


Figure 8: Positional Error from Original Position (0,75,0) for Sensor S_2 (with Gaussian noise)

Errors in determined coordinates with true Euclidean distance i.e. without Gaussian noise are almost negligible; this negligible error in turn validates the proposed mathematical model for coordinates determination with a single beacon. Besides, considering the physical sizes of the sensors deployed underwater, Figure 8 and 9 show errors (with Gaussian noise) are within acceptable range considering for a 200m water column.

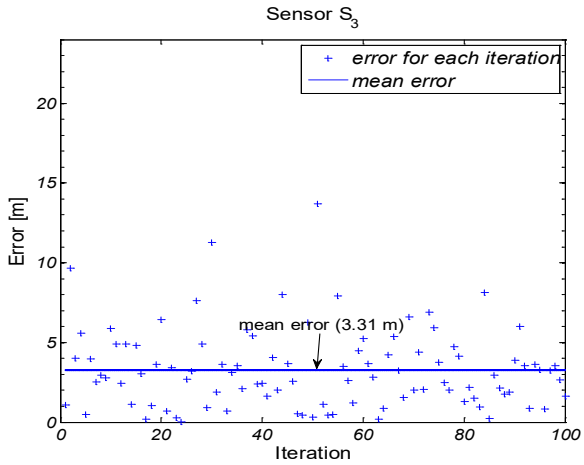


Figure 9: Positional Error from Original Position (80,40,0) for Sensor S_3 (with Gaussian noise)

For sensors with dimension 4.5x2x1.5cm, positional error 0.01-3.81cm is quite outstanding. As simulation in previous sections validates the model with the negligible error considering Euclidean distances, it is now conspicuous that the precision of the distances measured from the beacon to the deployed sensors produce lesser positional errors.

3.2. Experimental Results

A stringent experiment has been conducted in the terrestrial environments i.e. in the lab with compatible ultrasonic sensors to prove the proposed mathematical model for coordinates determination with a single beacon using Cayley-Menger determinant. It also shows that Cayley-Menger determinant could be used for non-parallel state situation with a little adjustment of the proposed model. Flight time i.e., signal's propagation delay of the acoustic signal is used to determine the distances between the surfaced beacon and deployed submerged sensors. A beacon at the ceiling and three other sensors on the table top have been kept to achieve that same scenario as the problem domain described in previous section.

Two different scenarios with four individual tests have been conducted to validate the model. Firstly, in scenario 1 positions of the sensors are S_1, S_2 and S_3 are (0,0,0), (0,20,0), (30,15,0) respectively. Keeping the origin of the Cartesian system on sensor S_1, positional error for sensors S_2 and S_3 are within 0.2 and 4cm range. Accuracy in distance measurements with the ultrasonic sensor generates positional error: 0.17cm, whereas in extreme case it is 3.85cm as depicted in Table 4. Secondly, in scenario 1 positions of the sensors are S_1, S_2 and S_3 are (0,0,0), (0,25,0), (35,10,0) respectively. Keeping the origin of the Cartesian system on sensor S_1, positional error for sensor S_2 and S_3 are within 0.5 and 6cm range. Accuracy in distance measurements with the ultrasonic sensors generates positional error: 0.17cm, whereas in extreme case it is 3.85cm as illustrated in Table 5.

The ultrasonic sensors used in the experiment have 12^0 sentry angle; this capacity limits the movement of the surfaced sensor. Besides the built-in Arduino *microseconds To Centimeter* function needed to be changed to 'double' to acquire more precise timing as the experiment is taken place where the maximum distances between beacon and deployed sensors are less than 80cm range.

One other challenge we faced while experimenting with Arduino board is to process the generated ultrasonic pulses that are received by the sensors to calculate inter distances. As in [14], the '*pulseIn*' function usually takes more than 20ms to process the received pulse, whereas it takes only 2.32ms to travel 80cm (approximate max distance for the experimental domain); as a result by the time it finishes pulse processing for the nearest sensor from the beacon and starts processing pulses for other sensors, it is then too late for the pulses to be on the flight. To mitigate this problem we had to generate two other 10µs pulses for rest of the two sensors in 50ms interval. So within around 100ms all three pulses are generated, this fraction of a second will have no effect on the stationary sensor nodes scenario; however, will have negligible effect in cases of mobility.

4. Conclusions

Persistent and accurate positioning is indispensable in various arenas due to necessity as well as research. As specific applications demand more accuracy and convenience, research pushes the boundary to fit and meet the demand in its own course. So, a plethora of localization algorithm has been proposed. This paper illustrates and analyzes a pragmatic approach of localization where a single node can determine deployed submerged nodes using Cayley-Menger determinant; it also portraits associated difficulties of using Cayley-Menger in non-parallel state situations in real time as the original model was designed for parallel states. In nature, parallel state situation is very rare where the plane of the beacon node and the plane of deployed sensors would be parallel. Hence, the necessity of having a pragmatic solution was conspicuous.

A solvable configuration of the domain and its associated model has been simulated to validate as well as to fathom the degree of errors. Once the distances between beacon and the deployed submerged sensors are considered to be true Euclidean, it generates negligible errors. It is also shown that the generated errors are because of erroneous distance determination; so the accuracy of coordinates solely depends on the preciseness of distance measurement method, not the model. This paper also delineates how the model could be used for non-parallel state situation with a little adjustment. Simulation results with Gaussian noise in distance measurements suggest that 0.62-3.67m error for a 200m water column is outstanding as sizes of deployed sensors or AUVs could reach few meters in length. However, mobility of the deployed sensors was not considered in this paper and left for future exploration. On the other hand, mobility of the beacon and span has very limited effect on the coordinates determination of the sensors.

To validate the simulation, a stringent experiment has been conducted in terrestrial environment. So called off-the-shelf ultrasonic sensors have been used imitating the simulated configuration. In the experiment, it has been shown that distance between beacon and deployed sensor can be calculated using acoustic signals, whereas clock synchronization could be performed using electrical signals. Mobility of the sensors was not considered in the experiment as well; however, multiple scenarios

Experimental Scenario 1: Original coordinates of sensors:

S_1: (0,0,0); S_2: (0,20,0); S_3: (30,15,0)

Table 4: Coordinates of the Sensors S_2 and S_3 according to Scenario 1

Trial 1

readings	S_1	S_2	S_3
R_1	54.14	52.31	50.36
R_2	57.96	52.65	51.10
R_3	51.32	50.56	52.87
R_4	61.11	52.49	55.71
R_5	59.90	58.57	48.99
R_6	60.14	55.04	50.17

Calculated Coordinates:
S_1: (0,0,0)
S_2: (0,19.83,0)
S_3: (29.58,14.94,0)

Trial 2

readings	S_1	S_2	S_3
R_1	62.57	57.27	52.20
R_2	60.31	54.78	53.17
R_3	53.39	52.60	55.01
R_4	63.58	54.62	57.96
R_5	62.32	60.94	50.98
R_6	56.34	54.42	52.40

Calculated Coordinates:
S_1: (0,0,0)
S_2: (0,20.55,0)
S_3: (30.88,15.45,0)

Trial 3

readings	S_1	S_2	S_3
R_1	60.76	55.20	53.57
R_2	56.76	54.83	52.79
R_3	53.80	53.00	55.43
R_4	64.06	55.03	58.40
R_5	62.79	61.41	51.36
R_6	63.04	57.71	52.60

Calculated Coordinates:
S_1: (0,0,0)
S_2: (0,20.73,0)
S_3: (30.93,15.42,0)

Trial 4

readings	S_1	S_2	S_3
R_1	61.82	53.11	56.36
R_2	58.64	53.27	51.70
R_3	51.92	51.15	53.49
R_4	54.78	52.92	50.95
R_5	60.60	59.26	49.57
R_6	60.84	55.69	50.76

Calculated Coordinates:
S_1: (0,0,0)
S_2: (0,20.01,0)
S_3: (29.94,15.01,0)

Experimental Scenario 2: Original coordinates of sensors:

S_1: (0,0,0); S_2: (0,25,0); S_3: (35,10,0)

Table 5: Coordinates of the Sensors S_2 and S_3 according to Scenario 2

Trial 1

readings	S_1	S_2	S_3
R_1	65.34	62.51	53.19
R_2	54.94	58.15	56.49
R_3	55.90	54.13	58.85
R_4	58.63	57.91	54.44
R_5	62.95	60.46	53.42
R_6	60.32	55.26	57.25

Calculated Coordinates:
S_1: (0,0,0)
S_2: (0,26.12,0)
S_3: (36.55,12.54,0)

Trial 2

readings	S_1	S_2	S_3
R_1	59.85	57.26	48.71
R_2	50.32	53.25	51.74
R_3	53.70	53.04	49.86
R_4	51.20	49.57	53.90
R_5	57.65	55.37	48.92
R_6	55.25	50.62	52.43

Calculated Coordinates:
S_1: (0,0,0)
S_2: (0,23.72,0)
S_3: (33.26,12.94,0)

Trial 3

readings	S_1	S_2	S_3
R_1	56.77	52.01	53.88
R_2	51.71	54.73	53.17
R_3	55.18	54.51	51.24
R_4	52.61	50.94	55.39
R_5	59.25	56.90	50.28
R_6	61.50	58.84	50.06

Calculated Coordinates:
S_1: (0,0,0)
S_2: (0,24.53,0)
S_3: (34.33,12.63,0)

Trial 4

readings	S_1	S_2	S_3
R_1	62.36	59.90	52.92
R_2	54.43	57.61	55.97
R_3	58.09	57.38	53.94
R_4	64.74	61.94	52.69
R_5	55.38	53.62	58.31
R_6	59.76	54.75	56.72

Calculated Coordinates:
S_1: (0,0,0)
S_2: (0,25.81,0)
S_3: (36.15,13.64,0)

and tests suggest the validity of the model. Acoustic signals are distressed by environmental variables like temperature and humidity as well as propagation speed is low compared to electrical signals. Besides, sensors that have been used in the experiment have limitations and constraints in signal processing; result indicates 0.01-3.81cm positional error for sensors with dimension 4.5x2x1.5cm is within acceptable range.

References

[1] J. H. Cui, J. Kong, M. Gerla, and S. Zhou, "The challenges of building mobile underwater wireless networks for aquatic applications," *Network, IEEE,* vol. 20, pp. 12-18, 2006.

[2] L. Hu, B. Liu, K. Zhao, X. Meng, and F. Wang, "Research and Implementation of the Localization Algorithm Based on RSSI Technology," *Journal of Networks,* vol. 9, pp. 3135-3142, 2014.

[3] P. Duff and H. Muller, "Autocalibration algorithm for ultrasonic location systems," in *Seventh IEEE International Symposium on Wearable Computers, 2003. Proceedings*, pp. 62-68, 2003.

[4] E. Olson, J. Leonard, and S. Teller, "Robust range-only beacon localization," in *Autonomous Underwater Vehicles, IEEE/OES*, pp. 66-75, 2004.

[5] J. Guevara, A. R. Jimenez, A. S. Morse, J. Fang, J. C. Prieto, and F. Seco, "Auto-localization in Local Positioning Systems: A closed-form range-only solution," in *IEEE International Symposium on Industrial Electronics (ISIE)*, pp. 2834-2840, 2010.

[6] J. C. Prieto, A. R. Jiménez, J. Guevara, J. L. Ealo, F. Seco, J. O. Roa, and F. Ramos, "Performance evaluation of 3D-LOCUS advanced acoustic LPS," *IEEE Transactions on Instrumentation and Measurement,* vol. 58, pp. 2385-2395, 2009.

[7] A. Rahman and V. Muthukkumarasamy, "Embracing localization inaccuracy with a single beacon", *Journal ofAdvanced Computer Science and Application (IJACSA)*, Vol. 10 No. 12, 2019.

[8] A. Rahman and V. Muthukkumarasamy, , "Localization of Submerged Sensors with a Single Beacon for Non-Parallel Planes State", in *10th International conference on Ubiquitous and Future Networks (ICUFN),* pp. 525-530, 2018.

[9] A. Rahman, V. Muthukkumarasamy, and E. Sithirasenan, "Coordinates Determination of Submerged Sensors Using Cayley-Menger Determinant," in *Distributed Computing in Sensor Systems (DCOSS), IEEE*, pp. 466-471, 2013.

[10] A. Rahman, V. Muthukkumarasamy, and E. Sithirasenan, "Localization of Submerged Sensors Using Radio and Acoustic Signals with Single Beacon," in *Ad-hoc, Mobile, and Wireless Network. LNCS.* vol. 7960, J. Cichoń, M. Gębala, and M. Klonowski, Eds., ed: Springer Berlin Heidelberg, pp. 293-304, 2013.

[11] I. Vasilescu, K. Kotay, D. Rus, M. Dunbabin, and P. Corke, "Data collection, storage, and retrieval with an underwater sensor network," in *Proceedings of the 3rd international conference on Embedded networked sensor systems*, pp. 154-165, 2005.

[12] J. Guevara, A. Jiménez, J. Prieto, and F. Seco, "Auto-localization algorithm for local positioning systems," *Ad Hoc Networks,* vol. 10, pp. 1090-1100, 2012.

[13] A. Rahman, V. Muthukkumarasamy, and X. Wu, "Coordinates and Bearing of Submerged Sensors Using a Single Mobile Beacon (CSMB)", *Journal of Networks*, Vol. 10 No. 8, 2015.

[14] P. Xie, J. H. Cui, and L. Lao, "VBF: vector-based forwarding protocol for underwater sensor networks," *Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems,* pp. 1216-1221, 2006.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 151-158 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

SiC-FET Gas Sensor for Detecting Sub-ppm Gas Concentrations

Yoshitaka Sasago[*, 1], Hitoshi Nakamura[1], Takahiro Odaka[1], Atsushi Isobe[1], Shigenobu Komatsu[1], Yohei Nakamura[1], Taizo Yamawaki[1], Chiko Yorita[2], Nobuyuki Ushifusa[2], Kohei Yoshikawa[3], Kazuo Ono[1], Yumiko Anzai[4], Shuntaro Machida[1], Masaharu Kinoshita[1], Koji Fujisaki[1], Kenji Okishiro[5], Yuta Sugiyama[6]

[1]*Center for Technology Innovation – Electronics, Hitachi, Ltd., Tokyo, 185-8601, Japan*

[2]*Center for Technology Innovation – Production Engineering, Hitachi, Ltd., Kanagawa, 244-0817, Japan*

[3]*Center for Technology Innovation – Materials, Hitachi, Ltd., Ibaraki, 319-1292, Japan*

[4]*Center for Technology Innovation – Healthcare, Hitachi, Ltd., Tokyo, 185-8601, Japan*

[5]*Global Research & Innovative Technology Center GRIT, Hitachi Metals, Ltd., Saitama, 360-8577, Japan*

[6]*Cable Materials Company, Hitachi Metals, Ltd., Ibaraki 319–1418, Japan*

ARTICLE INFO

Article history:
Received: 30 September, 2019
Accepted: 23 December, 2019
Online: 22 January, 2020

Keywords:
Capacitor
Platinum
YSZ
Nickel oxide
Nitrogen oxides (NOx)
Nitrogen monoxide (NO)
Oxygen (O_2)
Sulfur oxide (SO_2)
Carbon monoxide (CO)
Ammonia (NH_3)

ABSTRACT

This paper describes a silicon carbide-field effect transistor (SiC-FET) gas sensor that is able to detect NO, O_2, NH_3, CO, and SO_2. The gate of the sensor FET is a gas detection layer that consists of yttria-stabilized zirconia, nickel oxide, and platinum. The threshold voltages of the sensor depend on the target gas concentration, measurement temperature, and sensor gate materials. The device on SiC substrate can measure gas at temperatures up to 800°C. The gate material dependence and the wide-range temperature dependence of the sensitivity make it possible to detect the concentration of each component in a mixture of gases. In addition, a gate stack composed of composite materials improves the sensitivity of the FET sensor by helping to maintain the sensor's heat resistance and the reproducibility of its measurements.

1. Introduction

In particular, pollution in the atmosphere directly affects human health. Here, nitrogen oxides (NOx) are the most harmful components of automobile exhaust gas, and sensitive gas sensing is required for detecting them. In Ref. [1], we developed an SiC-FET-type NOx sensor for high-temperature exhaust gas with a sub-ppm nitrogen monoxide (NO) detection capability.

Limiting-current-type [2], mixed-potential-type [3], and resistive-type [4] sensors have been developed for NOx detection. The limiting-current type detects the electric current generated by oxygen ions from NO when a voltage is applied, and the oxygen-ion current depends on the NO concentration. This sensor can detect NO in a wide range of concentration, and it has been used in cars with diesel engines. The mixed-potential type detects voltages produced by NOx. The resistive type detects current flowing through gas-detection materials under an applied voltage. Compared with the limiting-current and resistive types, the SiC-FET sensor does not need a current to flow through the gas detection materials, because the current flows on the surface of the SiC. Furthermore, since the SiC-FET amplifies the gas-sensing signals, it is more sensitive than the mixed-potential type [3].

[*]Yoshitaka Sasago, 1-280 Higashi-Koigakubo, Kokubunji, Tokyo 185–8601, Japan, +81-42-323-1111 & yoshitaka.sasago.tx@hitachi.com

The SiC-FET NOx sensor can detect NO in the mixture of gases in exhaust fumes when it is equipped with an ion-pumping device for preprocessing. The ion-pumping device can remove disturbance gases, and thereby, it enhances the selectivity of the sensor for NO gas [2, 5]. Moreover, without the on-chip ion-pumping device, the sensor can detect several gases such as oxygen (O_2), sulfur oxide (SO_2), carbon monoxide (CO), and ammonia (NH_3). These gases are also included in the exhaust from automobile engines. Instead of improving the sensor's selectivity by removing disturbance gases, we propose that it can be enhanced by using the temperature dependence of gas sensitivity. The high-temperature immunity of SiC-FETs suggests that they can be used for gas-sensing operation over a wide range of temperatures, and the sensitivity to each gas component of these devices varies depending on the sensing temperature.

In this study, we examined the temperature dependence of the gas responses of the SiC-FET sensor. We found that it can sensitively detect several different gases at high temperature. We conclude that a low-cost gas sensor with high sensitivity can be realized by taking advantage of semiconductor products.

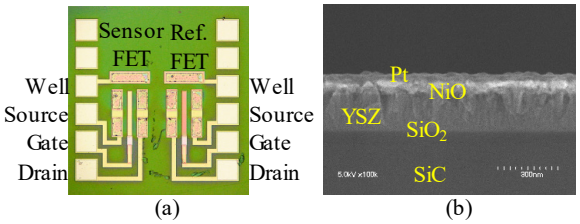


Figure 1: SiC-FET-type NOx sensors [1]: (a) photograph of SiC chip (2 mm × 2 mm), (b) gate layer of sensor FET.

2. Device Structure and Operation Principle

As shown in Fig. 1(a), the SiC-FET NOx sensor chip measuring 2 × 2 mm^2 includes a sensor FET, a reference FET, and platinum electrodes for the well, source, gate, and drain terminals on the edges on both sides [1]. Both FETs have a gate layer stack of yttria-stabilized zirconia (YSZ), nickel oxide (NiO), and platinum (Pt) on a gate oxide layer (Fig. 1 (b)). The platinum layer of the sensor FET is exposed to the atmosphere, while that of the reference FET is covered with dielectric films. The composition of YSZ is $(ZrO_2)_{0.97}(Y_2O_3)_{0.03}$. The process flow of the sensor chip is explained in Ref. [1]. Since NO, O_2, SO_2, CO, and NH_3 can only be sensed at high temperature, a heater should be placed near the sensor. The heater can be included in a module along with the sensor.

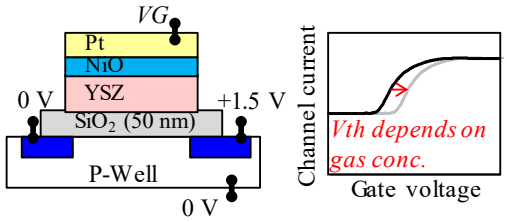


Figure 2: Device structure and operation principle of FET-type sensor [1].

The operation principle of the SiC-FET gas sensor is described in Ref. [1]. As shown in Fig. 2, the threshold voltage (*Vth*) of the sensor FET shifts by increasing the target gas concentration. For example, when the target gas is NO, *Vth* shifts in the positive direction as the concentration increases. However, *Vth* shifts in the negative direction for NH_3. The target gas concentration can be calculated from the *Vth* shift. A SiC-FET gas sensor was studied for hydrogen detection [6] and ammonia detection [7], but its gate structure was different from our device. The gate material can facilitate detection of target gases. The FET- and capacitor-type gas sensors are classified as work function-type sensors [8].

3. Gas Detection Material

FET sensors have been studied for detection of hydrogen [6, 9–10], ammonia [7], oxygen [11], NH_3 and H_2S [12], etc. The gate layers of these sensors consist of detection materials, which determine the gas detection characteristics of the sensors. Here, we examined the oxygen gas response of the sensor because our primary targets are oxygen-related gases. To select suitable gate-layer materials, we compared the oxygen gas response of the SiC-FET gas sensors with those of Si-FET gas sensors.

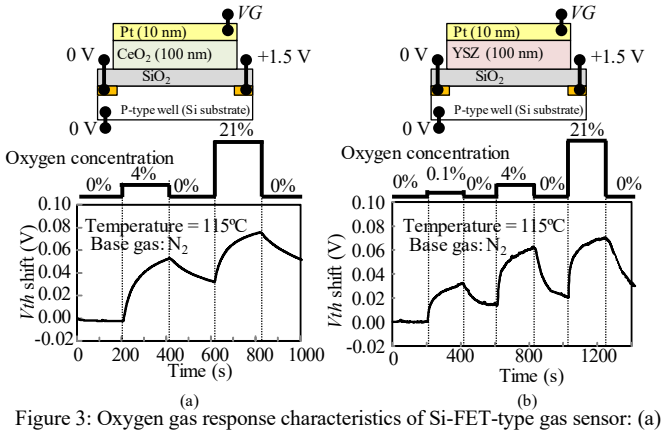


Figure 3: Oxygen gas response characteristics of Si-FET-type gas sensor: (a) platinum-ceria gate, (b) platinum-YSZ gate.

The oxygen response characteristics of the sensors with a platinum-YSZ gate and a platinum-ceria gate are compared in Fig. 3. The platinum, ceria, and YSZ films were deposited by sputtering. YSZ and ceria are oxygen ion conductors [13, 14], and platinum is a commonly used electrode for gas sensors [9–12]. Both sensors responded to oxygen gas; *Vth* shifted in the positive direction at 115°C, and the shifts depended on the oxygen concentration.

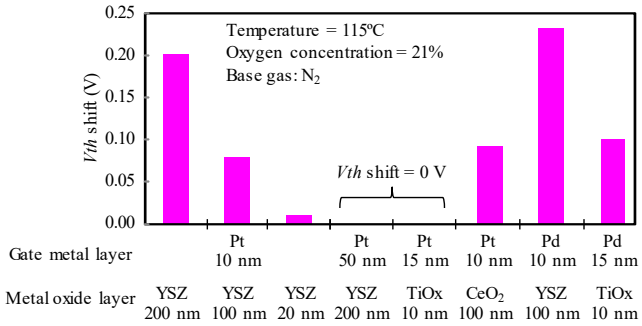


Figure 4: *Vth* shifts of Si-FET-type gas sensors with different gate layer materials.

The *Vth* shifts of different gate-layer materials are summarized in Fig. 4 for 21% oxygen at 115°C. The shifts depended on the thickness of the platinum layer, which is consistent with the results in Ref. [11]. The sensor with 10-nm-thick platinum responded to

oxygen gas, but the one with 50-nm-thick platinum did not. The thickness of YSZ also affected the *Vth* shift. The sensor with 10-nm-thick platinum and 200-nm-thick YSZ responded to oxygen gas. However, *Vth* shift decreased with decreasing YSZ thickness, and the sensor with 20-nm-thick YSZ hardly responded to oxygen gas. Platinum-titania gate FET responded to hydrogen gas, but not oxygen gas [9, 10]. The palladium-titania gate sensor responded to oxygen gas, but the platinum-titania gate sensor did not. The response of the sensor with 10-nm-thick palladium and 100-nm-thick YSZ was the largest of the test samples.

Among the oxide and metal materials in Fig. 4, platinum and YSZ are stable in terms of their chemistry and mechanics. Ceria is reduced in hydrogen atmosphere [14], and palladium is mechanically fragile because its volume expansion and contraction are large under a change in the hydrogen concentration. For these reasons, we selected platinum and YSZ for our SiC-FET gas sensor. The problem is that a thin-film platinum layer does not resist high temperatures over 500°C, while a sensor with a thick platinum layer (50 nm) does not respond to oxygen at 115°C, as shown in Fig. 4. However, a sensor thick platinum layer can do so at higher temperatures, as described below. We decided to use a 100-nm-thick platinum layer.

Although the sensor reported in Ref. [3, 15] is not an FET type, in that study it was shown that a catalytic metal oxide layer helps to enhance sensitivity for NOx [3, 15]. NiO and WO_3 are promising candidates for the catalytic layer, and one made of NiO shows excellent catalytic characteristics. Therefore, we decided to use a NiO layer in our sensor. In particular, NiO layers with thicknesses of 10 nm and 50 nm were formed between the platinum and YSZ layers. Moreover, composite material layers consisting of Pt and NiO were used instead of stacks of NiO and Pt.

The composition of the YSZ film was measured by X-ray photoelectron spectroscopy (XPS). It was the same as that of the sputtering target ($(ZrO_2)_{0.97}(Y_2O_3)_{0.03}$). No segregation of yttria was observed.

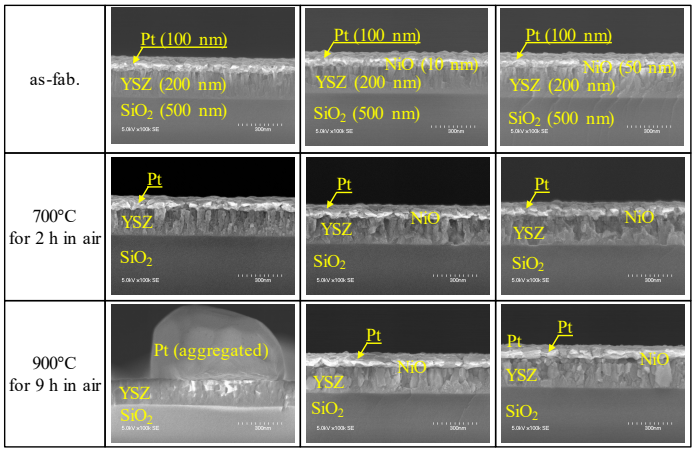


Figure 5: Heat resistance of sensor gates with 100-nm thick platinum. The photographs in the left and central columns are from Ref. [1].

The heat resistance of the gate stacks was examined by taking SEM micrographs. The results for the Pt-YSZ stacks and Pt-NiO-YSZ stacks are shown in Fig. 5. The platinum layer without the NiO layer became aggregated and disconnected at 900°C, but those with the 10-nm-thick and 50-nm-thick NiO layers remained continuous. The results for the stacks with composite material layers are shown in Fig. 6. The composite material layers also showed heat resistance up to 900°C. These results indicate that gate stacks with NiO could possibly achieve gas sensing at temperatures up to around 900°C.

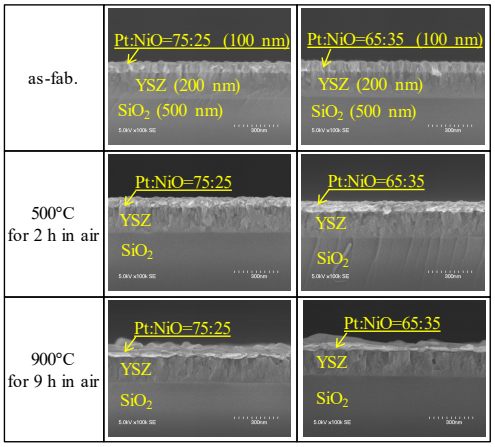


Figure 6: Heat resistance of sensor gates with Pt-NiO composite materials. The photographs in the left columns are from Ref. [1].

4. Gas-Sensing Characteristics

By virtue of the large band gap of SiC, SiC FETs are capable of on- and off-switching at over 500°C; thus, the *Vth* shift of the SiC-FET gas sensor can be detected [1]. On the other hand, switching at 500°C is impossible for Si FETs.

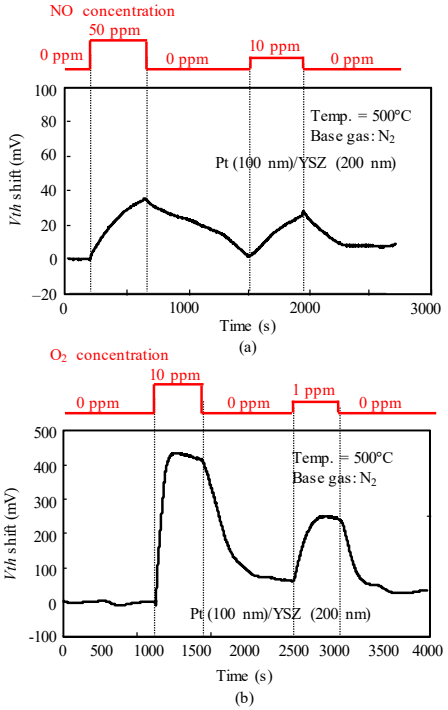


Figure 7: Gas response characteristics of the SiC-FET-type gas sensor with Pt-YSZ gate: (a) NO response, (b) O_2 response.

The time dependence of *Vth* was measured for the SiC-sensor FET. As Figs. 7 and 8 show, *Vth* depended on the concentrations of NO and O_2. Moreover, as shown in Fig. 7, the sensor with the Pt-YSZ gate responded to both NO and O_2, but its sensitivity to O_2

was higher. It clearly responded to 1 ppm O_2, with a *Vth* shift of about 250 mV, and it responded to 10 ppm NO with a *Vth* shift of about 25 mV. As well, the Pt-NiO-YSZ gate FET responded to both NO and O_2. The *Vth* shift for 1 ppm O_2 was about 200 mV, and that for 10 ppm NO was about 75 mV. The selectivity for NO and O_2 was thus modified by inclusion of the NiO layer.

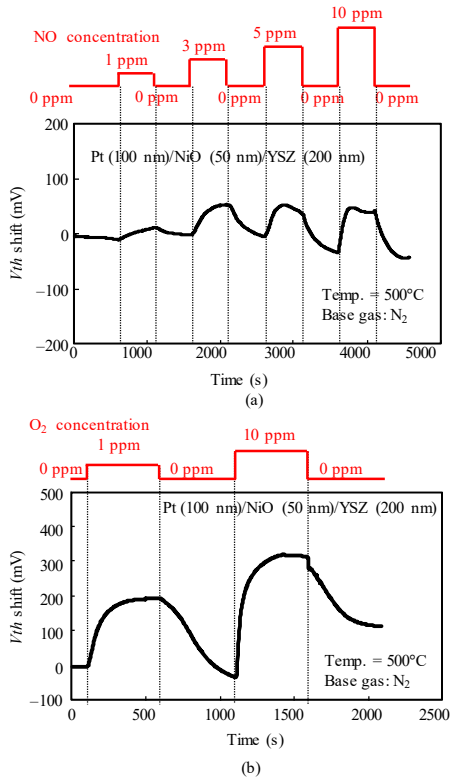


Figure 8: Gas response characteristics of the SiC-FET-type gas sensor with Pt-NiO-YSZ gate [1]: (a) NO response, (b) O_2 response.

We fabricated SiC-FET sensors on SiC wafers and measured the gate-length dependence of the FET characteristics of seven samples in air. In particular, *Vth* of the FET decreased smoothly as the gate length decreased; therefore, the sample variation was small. However, the gas response characteristics were measured in only one sample with a gate length of 100 nm for the sensor with the Pt-YSZ gate and the one with the Pt-NiO-YSZ gate. The sample-to-sample variation of gas responses thus remains an issue for future study.

Because of the sample variation problem, we decided to measure the responses of MOS-capacitor-type sensors with N-type SiC substrates instead for NO, O_2, SO_2, CO, and NH_3. While MOS capacitors are simpler than FETs, a change in the target gas concentration produces *Vth* shifts in these sensors like in FET sensors. Both FET- and capacitor-type gas sensors are classified as work function-type sensors [8].

As shown in Fig. 9, *Vth* of the Pt-YSZ gate capacitor increased as the concentration of NO gas increased, like the response of the FET sensor shown in Fig. 7(a). We measured the responses of a Pt-NiO-YSZ gate capacitor with 10-nm-thick NiO for NO gas at 700 and 800°C (Fig. 10) and those of a gate capacitor with 50-nm-thick NiO layer for NO gas at 600 and 800°C (Fig. 11). As Fig. 11 (a) shows, the sensor with 50-nm-thick NO layer responded to 0.5-ppm NO at 600°C, and the shift for 50 ppm NO was larger than 300 mV. Comparing Fig. 9 with Fig. 11 (a) suggests that the NiO layer appears to improve the NO sensitivity of the sensor. We also measured the NO responses of the capacitor-type sensor at 400°C, but the sensor hardly responded.

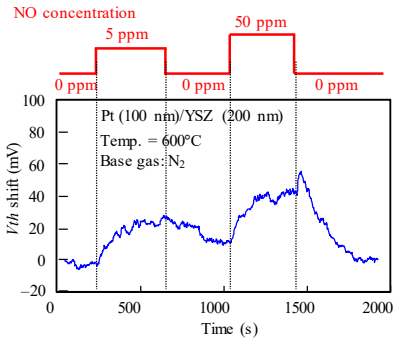


Figure 9: NO response characteristics of SiC-capacitor-type gas sensor with Pt-YSZ gate.

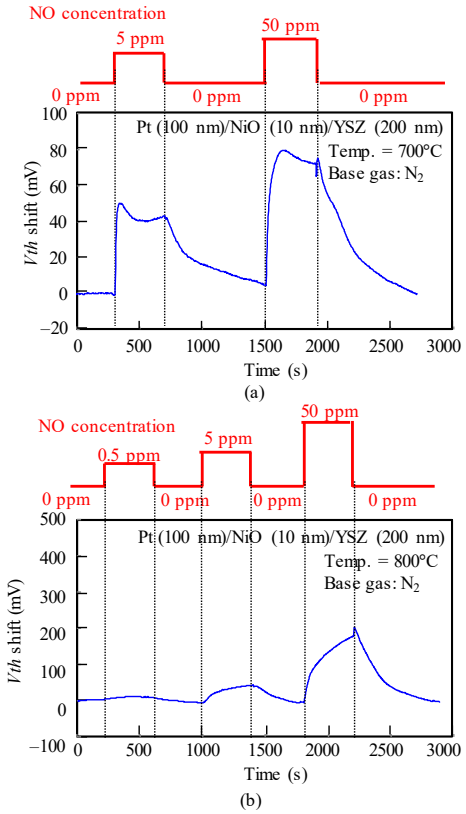


Figure 10: NO response characteristics of SiC-capacitor-type gas sensor with Pt-NiO (10 nm)-YSZ gate: (a) 700°C, (b) 800°C.

We measured the sensor responses of a Pt-YSZ gate capacitor for O_2 gas at 400 and 600°C and those of a Pt-NiO-YSZ gate capacitor with 10-nm-thick NiO for O_2 gas at 400, 600, and 800°C (Figs. 12 and 13). The sensors responded to an O_2 concentration of 0.1 ppm even at 400°C. The O_2 response could be detected at 400°C, while the NO response could not, because the catalytic reaction of the gate layers for O_2 is more active than for NO at low temperatures.

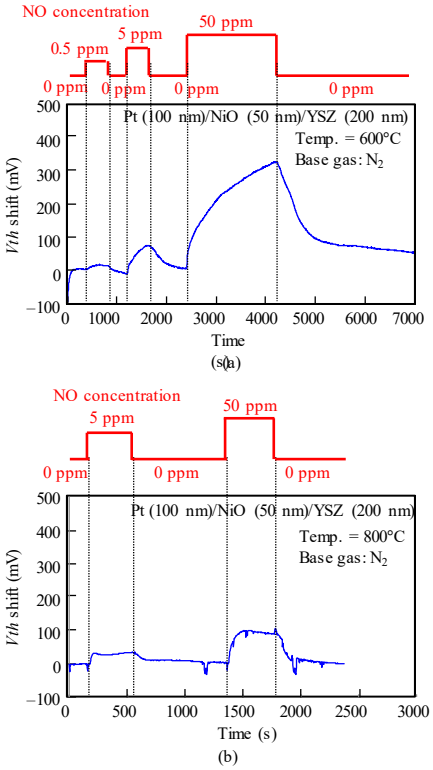


Figure 11: NO response characteristics of SiC-capacitor-type gas sensor with Pt-NiO (50 nm)-YSZ gate: (a) 600°C, (b) 800°C.

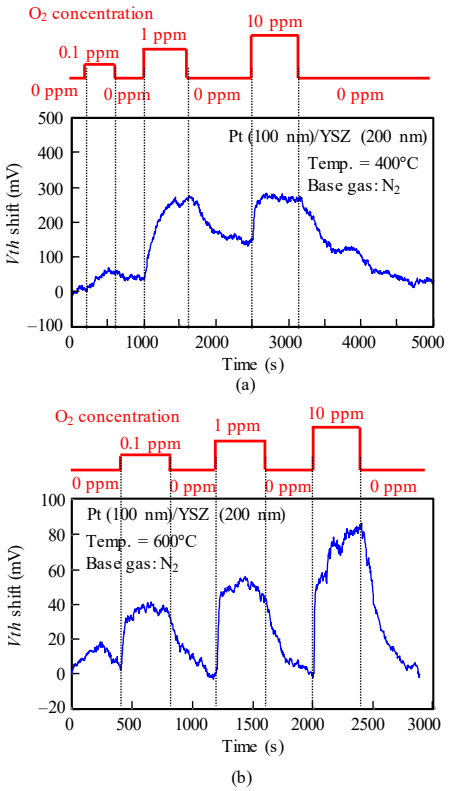


Figure 12: O_2 response characteristics of SiC-capacitor-type gas sensor with Pt-YSZ gate: (a) 400°C, (b) 600°C.

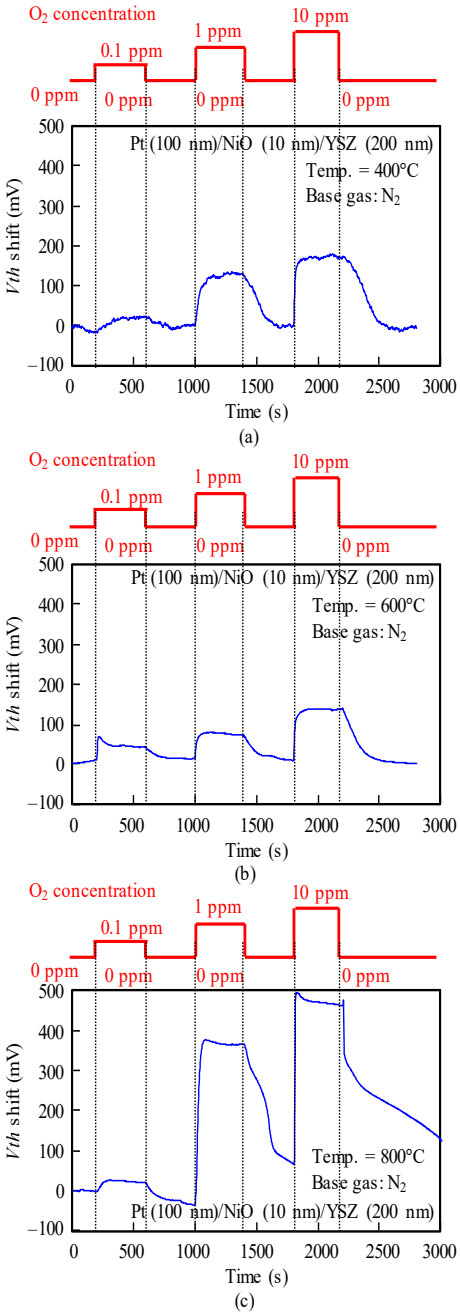


Figure 13: O_2 response characteristics of SiC-capacitor-type gas sensor with Pt-NiO (10 nm)-YSZ gate: (a) 400°C, (b) 600°C, (c) 800°C.

The responses of a Pt-YSZ gate capacitor were measured for NH_3 gas at 400, 600, and 700°C (Fig. 14), and those of a Pt-NiO-YSZ gate capacitor with 10-nm-thick NiO were measured for NH_3 gas at 400, 600, 700, and 800°C (Figs. 15 and 16). Although *Vth* shifted in the positive direction for NO and O_2 (Figs. 7–13), it shifted in the negative direction for NH_3. Note that a negative shift in *Vth* was previously observed in FET-type hydrogen sensors [9, 10]. Both sensors responded to NH_3 of 29 ppm concentration at 400°C. At 700°C, the response of the Pt-YSZ gate capacitor dropped rapidly, but that of a Pt-NiO-YSZ gate capacitor with 10-nm NiO decreased only slightly.

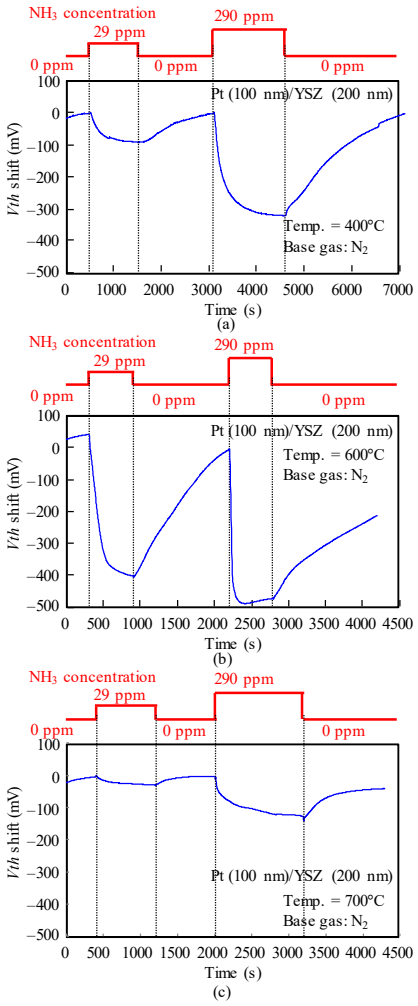


Figure 14: NH_3 response characteristics of SiC-capacitor-type gas sensor with Pt-YSZ gate: (a) 400°C, (b) 600°C, (c) 700°C.

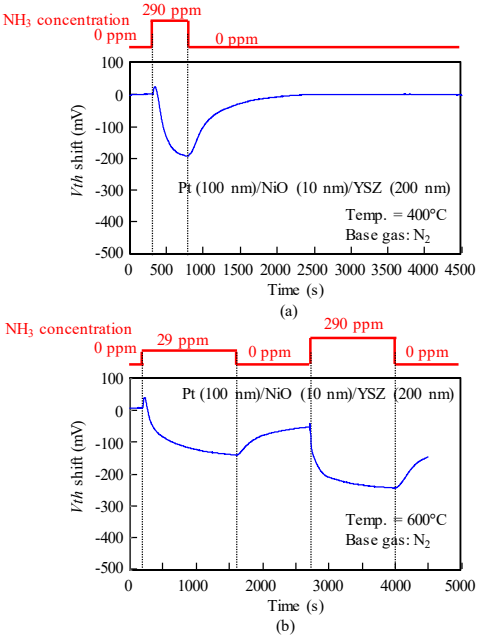


Figure 15: NH_3 response characteristics of SiC-capacitor-type gas sensor with Pt-NiO (10 nm)-YSZ gate: (a) 400°C, (b) 600°C..

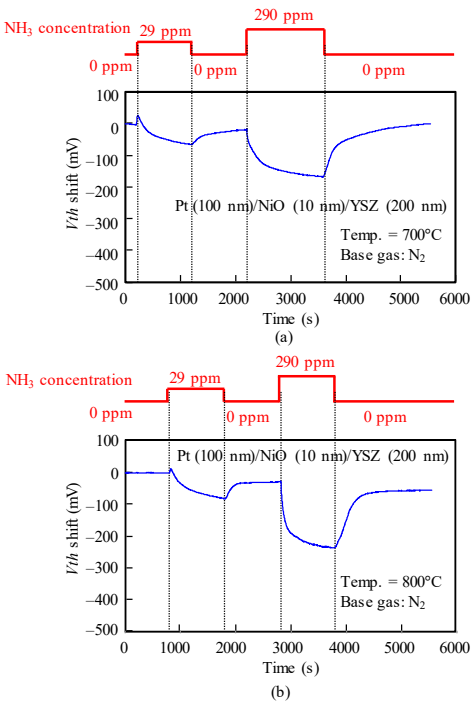


Figure 16: NH_3 response characteristics of SiC-capacitor-type gas sensor with Pt-NiO (10 nm)-YSZ gate: (a) 700°C, (b) 800°C.

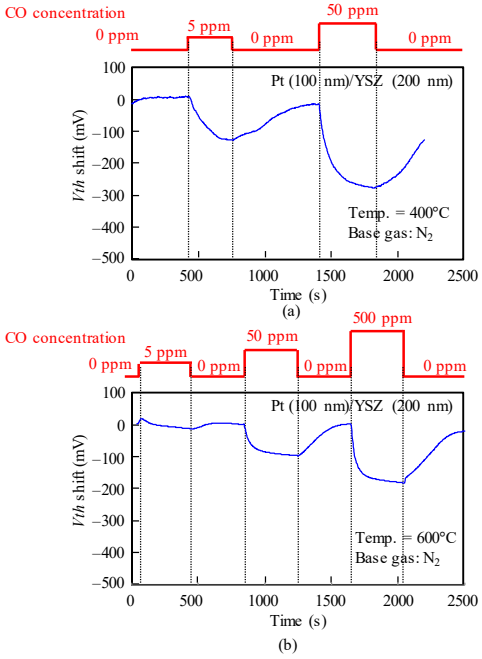


Figure 17: CO response characteristics of SiC-capacitor-type gas sensor with Pt-YSZ gate: (a) 400°C, (b) 600°C.

The sensor responses of the Pt-YSZ gate capacitor were measured for CO gas at 400 and 600°C (Fig. 17) and those of the Pt-NiO-YSZ gate capacitor with 10-nm-thick NiO were also measured for CO gas at 400, 600, and 700°C, as shown in Fig. 18. Similar to the responses for NH_3, *Vth* shifted in the negative direction for CO. The sensors responded to a CO concentration of a 5 ppm at 400°C. As the temperature increased, the sensor response of the capacitors decreased. The Pt-NiO-YSZ gate capacitor was measured at 800°C, but it did not respond to CO gas.

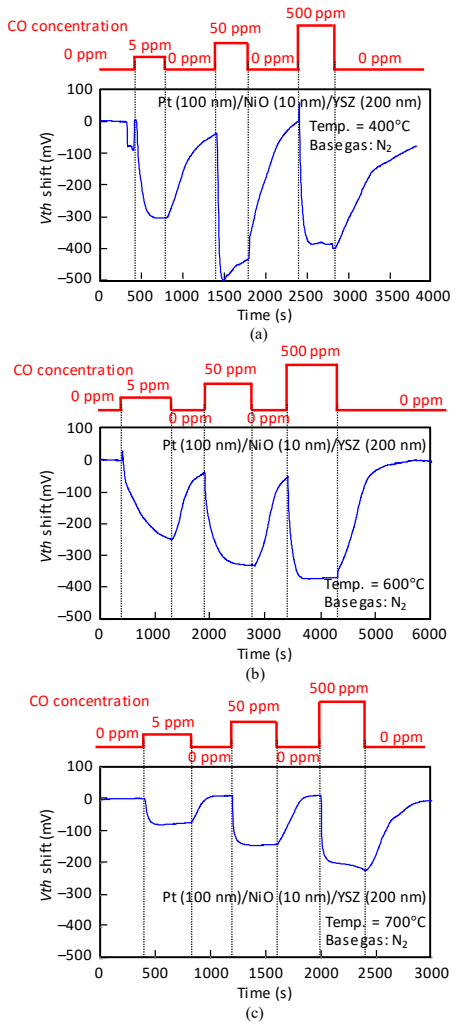


Figure 18: CO response characteristics of SiC-capacitor-type gas sensor with Pt-NiO (10 nm)-YSZ gate: (a) 400°C, (b) 600°C, (c) 700°C.

We measured the sensor responses of a Pt-NiO-YSZ gate capacitor with 50-nm-thick NiO for SO_2 gas at 600 and 800°C. As shown in Fig. 19, *Vth* shifted in the positive direction for SO_2 gas, as in the case of NO and O_2 gases. The sensor responded to a 5 ppm concentration of SO_2 at 600°C, and the response slightly decreased at 800°C.

The temperature dependence of the *Vth* shifts are summarized in Fig. 20 for NO (5 ppm), O_2 (1 ppm), NH_3 (29 ppm), CO (5 ppm), and SO_2 (5 ppm) gases. The results clearly show that the sensitivity depends on the measurement temperatures and the composition of the gate stack layer. Moreover, the measurement temperature can be controlled by the heater attached to the sensor module. Overall, the results indicate that measurement is possible over a range from 400 to 800°C, which is impossible for Si FET sensors.

The sensor responses of a Pt-YSZ gate capacitor and a Pt-NiO-YSZ gate capacitor were measured at room temperature, but no *Vth* shift was observed for NO (5 ppm), O_2 (1 ppm), NH_3 (29 ppm), CO (5 ppm), or SO_2 (5 ppm) gases. These results indicate our sensors have threshold temperatures for gas detection. Similar results are reported for NH_3 and H_2S in Ref. [12], in which the gas response was observed above 200°C for H_2S at a concentration of 400 ppm and above 110°C for NH_3 at a concentration of 500 ppm.

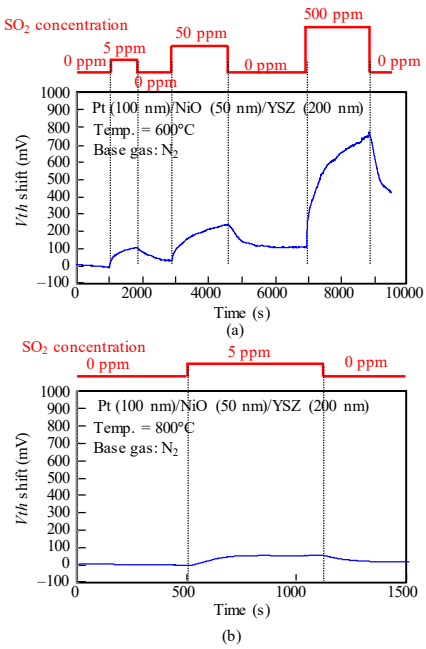


Figure 19: SO_2 response characteristics of SiC-capacitor-type gas sensor with Pt-NiO (50 nm)-YSZ gate: (a) 600°C, (b) 800°C.

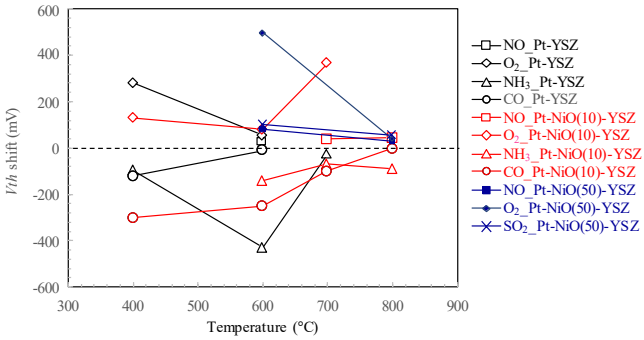


Figure 20: Temperature dependence of *Vth* shifts for NO (5 ppm), O_2 (1 ppm), NH_3 (29 ppm), CO (5 ppm), and SO_2 (5 ppm) gases.

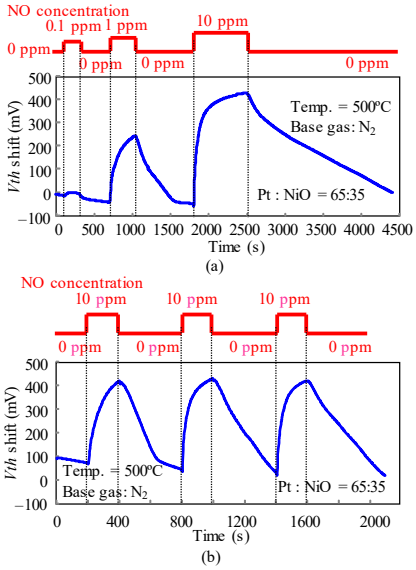


Figure 21: NO response characteristics of SiC-capacitor-type gas sensor with Pt-NiO-composite-material gate: (a) concentration dependence [1], (b) reproducibility.

To increase the sensitivity of the FET-type sensor, the three-phase boundary of the gate should be increased [16]. The nanoporous structure of the gate helps to improve the sensor sensitivity [9, 10]. The sensor with a NiO-Pt composite material gate showed NO detection with 0.1 ppm (Fig. 21 (a)). The reproducibility of the measurements was confirmed for 10 ppm NO, as shown in Fig. 21 (b). As mentioned above and as shown in Fig. 6, this sensor was very heat resistant.

5. Conclusion

We developed a SiC-FET-type gas sensors and SiC-capacitor-type gas sensors with a gate stack including YSZ, nickel oxide and platinum. These sensors are able to detect NO, O_2, NH_3, CO, and SO_2. To study the feasibility of the gas detection of each component in the gas mixture, we examined the temperature dependence of the gas response characteristics for each gas.

The nickel oxide layer enhances the heat resistance of the gate stack, and, thus, it extends the range of gas detection temperatures up to 800°C. The nickel oxide layer also helps enhance the sensitivity for NO detection. The sensors successfully detected NO at a concentration of 0.1 ppm. They were also able to respond to O_2 at a concentration of 0.1 ppm.

The gas response characteristics of capacitor-type sensors were measured or NO, O_2, NH_3, CO, and SO_2. The results show that *Vth* shifts in the positive direction for NO, O_2, and SO_2, but shifts in the negative direction for NH_3 and CO. The temperature dependence of the sensitivity was measured at room temperature and from 400 to 800°C. Gas responses were not observed at room temperature, and the sensitivity and the selectivity depended on the measurement temperature and the gate material of the sensors. This results indicate that the concentrations of each gas component can be determined using these dependences. In the future, we will study the sample variation of the gas responses characteristics and derive quantitative formulae for the concentration of each gas component.

In addition to the sensor with a gate stack of YSZ, NiO and Pt, the sensor with a NiO-Pt composite material gate was fabricated, and it was able to detect NO with a concentration of 0.1 ppm. This sensor was heat resistant, and its measurements were reproducible.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors gratefully acknowledge Dr. T. Sekiguchi and Dr. H. Kurata of Hitachi, Ltd. for all their encouragement. The authors also thank all the staff of the clean room of Hitachi's R&D group for the device fabrication. The authors thank Dr. R. Goto, Dr. H. Sano, Dr. T. Shimada, Mr. K. Kato, Mr. N. Ashizuka, and Mr. Y. Nozawa of Hitachi Metals, Ltd. for insightful discussion from the early stage of the research. The authors also thank Mr. Y. Nagai of TOMOE SHOKAI Co., LTD. for valuable assistance with the gas response experiments.

References

[1] Y. Sasago et al., "SiC-FET-type NOx Sensor for High-Temperature Exhaust Gas," 2018 IEEE International Electron Devices Meeting, 12.5.1–12.5.4.

[2] N. Kato, K. Nakagaki, and N. Ina, "Thick Film ZrO_2 NO_x Sensor," SAE Technical Paper 960334 (1996).

[3] N. Miura et al., "Improving NO_2 Sensitivity by Adding WO_3 during Processing of NiO Sensing-Electrode of Mixed-Potential-Type Zirconia-Based Sensor," J. Electrochem. Soc. 154, pp. 1246–1252 (2007).

[4] A. Groß et al., "Dual Mode NO_x Sensor: Measuring Both the Accumulated Amount and Instantaneous Level at Low Concentrations," Sensors 12, pp. 2831-2850 (2012).

[5] T. Inaba and T. Saji, "Low temperature operation of thin-film limiting-current type oxygen sensor using graded-composition layer electrodes," Sensors and Actuators B 129, pp. 874–880 (2008).

[6] R. Loloee et al., "Hydrogen monitoring for power plant applications using SiC sensors," Sensors and Actuators B 129, 200–210 (2008).

[7] H. Wingbrant et al., "Using a MISiC-FET Sensor for Detecting NH_3 in SCR Systems," IEEE Sensors Journal 5, pp. 1099–1105 (2005).

[8] T. Hübert et al., "Hydrogen sensors – A review," Sensors and Actuators B 157, pp. 329–352 (2011).

[9] T. Usagawa and Y. Kikuchi, "A Pt–Ti–O gate Si-metal-insulator-semiconductor field-effect transistor hydrogen gas sensor," J. Appl. Phys. 108, 074909 (2010).

[10] Y. Sasago et al., "FET-type hydrogen sensor with short response time and high drift immunity," Symp. on VLSI Technol. 2017, T106–T107.

[11] Y. Miyahara, K. Tsukada, and H. Miyagi, "Field-effect transistor using a solid electrolyte as a new oxygen sensor," J. Appl. Phys. 63, 2431 (1988).

[12] T. Usagawa and Y. Sasago, "Hydrogen-Related Gas Response from Pt-Ti-O Gate Si-MISFETs Gas Sensors," Proceedings of the 62nd Chemical Sensor Symposium 31, pp. 16-17 (2017).

[13] S. J. Skinner and J. A. Kilner, "Oxygen ion conductors," Materials Today 6, pp. 30–37 (2003).

[14] Y. Zhang et al., "Recent Progress on Advanced Materials for Solid-Oxide Fuel Cells Operating Below 500°C," Adv. Mater. 29, 1700132 (2017).

[15] A. Hemberga et al., "Effect of the thickness of reactively sputtered WO_3 submicron thin films used for NO_2 detection", Sensors and Actuators B 171, pp. 18-24 (2012).

[16] B. Wang et al., "Fabrication of Well-Ordered Three-Phase Boundary with Nanostructure Pore Array for Mixed Potential-Type Zirconia-Based NO_2 Sensor," ACS Appl. Mater. Interfaces, 8, pp. 16752–16760 (2016).

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 2, 154-165 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Transfer Learning and Fine Tuning in Breast Mammogram Abnormalities Classification on CBIS-DDSM Database

Lenin G. Falconi[*,1], Maria Pérez[1], Wilbert G. Aguilar[1], Aura Conci[2]

[1]*Escuela Politécnica Nacional, Facultad de Ingeniería de Sistemas, 170517, Ecuador*

[2]*Federal Fluminense University, VisualLab, Computer Institute, 24210-240, Brazil*

ARTICLE INFO

Article history:
Received: 15 January, 2020
Accepted: 18 February, 2020
Online: 11 March, 2020

Keywords:
Transfer Learning
Fine Tuning
Convolutional Neural Networks
Mammogram classification

ABSTRACT

Breast cancer has an important incidence in women mortality worldwide. Currently, mammography is considered the gold standard for breast abnormalities screening examinations, since it aids in the early detection and diagnosis of the illness. However, both identification of mass lesions and its malignancy classification is a challenging problem for artificial intelligence. In this work, we extend our previous research in mammogram classification, where we studied NasNet and MobileNet in transfer learning to train a breast abnormality malignancy classifier, and include models like: VGG, Resnet, Xception and Resnext. However, training deep learning models tends to overfit. This problem is also carried out in this work. Our results show that Fine Tuning achieves the best classifier performance in VGG16 with AUC value of 0.844 in the CBIS-DDSM dataset.

1 Introduction

This work is an extension of our work originally presented in IWSSIP 2019 [1] about mammogram abnormalities classification using *Transfer Learning* (TL) with Mobilenet [2] and Nasnet [3]. In this paper, we also address the classification problem of mammogram abnormalities using the CBIS-DDSM [4] dataset, but we extend the experimentation in transfer learning to other ImageNet pre-trained convolutional neural network (ConvNet) models like: Resnet, Resnext, Xception; to name a few. Finally, *Fine Tuning* (FT) is used in order to address the overfitting problem and improve previous results on the CBIS-DDSM dataset.

Despite the increase in understanding of breast cancer as a disease, it is still a major public health problem worldwide because of the incidence and mortality rates it presents [5]. According to the International Agency for Research on Cancer (IARC), this illness is the second most frequent form of cancer among women worldwide with 2,088,849 (11.6%) new cases and 626,679 (6.6%) of deaths [6]. The mammogram exam remains the gold standard for screening examination, mainly because it is the only screening test that has proven to reduce mortality [7]. However, mammography has some limitations like the variability of its sensitivity, which is inversely proportional to breast density, the false positive and negative rates, and the patient's exposure to radiation [7]. Other screening tests available are: ultrasound, magnetic resonance (MRI), tomosynthesis, and infrared thermography [8]-[9]; in most cases, the aforementioned screening tests are used as adjunct tests.

The mammogram exam diagnostic relies on the radiologist's experience for detection. However, 10% of all woman screened for cancer are called back for additional testing and just as little as 0.5% of them are diagnosed with breast cancer [10]. This shows that it is important to design CAD systems that aid specialists, and train new ones, in breast lesions detection. A "*classic*" CAD system is comprised of 5 main stages: image pre-processing, image segmentation or region of interest (ROI) definition, feature extraction and selection, classification, and performance evaluation [9, 11]. However, this model can be said to be in change due to advances in the field of machine learning, specifically in *Deep Learning*; which allows to automatically learn representations of data with multiple levels of abstraction through deep convolutional neural networks [12]. For instance, in the field of computer vision, the classification of natural images has shown an incredible increase in performance since 2012, when the AlexNet ConvNet model to classify natural images in 1000 categories presented in [13] achieved a 15.3% top 5 test error rate. As a matter of fact, the stages of feature extraction and classification can be solved directly by a ConvNet [14]. This reduces the need for feature hand engineering, which was tradition-

[*]Corresponding Author: Lenin G. Falconi, lenin.falconi@epn.edu.ec

ally used to create the feature vector, because a ConvNet is able to synthesize its own feature vector [15]. All of this confirms the point indicated in [11] that the development in both image techniques and computer science enhance the interpretation of medical images.

The success of ConvNets and deep learning in computer vision tasks such as image classification heavily relies on the number of examples used in training the model under the supervised learning paradigm [16]. Unfortunately, mammogram public datasets are not "deep enough". In this context, transfer learning and fine tuning are deep learning techniques that can aid the development of accurate enough classifiers by transferring knowledge from another domain where large datasets are available. One of the main problems when dealing with small number of examples in training is overfitting. Transfer Learning and Fine Tuning aid in overcoming this disadvantage of working with ConvNets.

In this work, we aim to classify region of interest images from mass tumors of the CBIS-DDSM [4] dataset. We extend our previous experimentation presented in [1] by using different ConvNets models and also the Fine Tuning technique in order to increase the performance of the classifier of breast mammogram abnormalities in benign and malignant. Our research results indicate that Fine Tuning is able to train an accurate classifier and overcome overfitting. Also, we have included the ROC curve metric to measure the performance of the classifiers here studied.

The remainder of this paper is organized as follows: In Section 2, we perform a review of machine learning concepts related to the current research; specifically convolutional neural networks, transfer learning and fine tuning. Literature review of related works in the field is in Section 3. Our proposed experimental method, dataset, and model are presented in Section 4. Section 5 presents experimental data results. Finally, discussion and future works are presented in Section 6 and 7, respectively.

2 Deep Learning Background

2.1 Convolutional Neural Networks

Models based on the Convolutional Neural Networks (ConvNets) architecture have been able to achieve high accurate results in image classification and detection tasks in the ImageNet[17] dataset, under the supervised learning paradigm and back-propagation. That is the case of residual networks, proposed in [18], which achieved a 3.57% error rate in the ImageNet test set in 2015.

Traditional pattern recognition classifiers rely on a hand designed feature extractor that derives relevant information from the raw input data [16]. Thus, the feature extraction step aims to reduce the dimension of the data while characterizing the raw input data (image, sound, etc.) meaningfully so that a trainable classifier is able to categorize its feature vectors [16, 19]. However, the design of the feature extractor requires specialized knowledge about the data (hand-engineering) that, in some cases, could be unknown [19]. On the contrary, ConvNets eliminate the feature extraction process by absorbing it in their architecture [16]. As pointed out in [14], the structure of a ConvNet combines both: the feature extraction and classification steps in one single model that is trained on back-propagation; the feature extraction task is, therefore, learned from data in the first layers of the model, while the last full connecting layers constitute the classifier task. The LeNet-5 ConvNet, proposed in [15] to solve the handwritten classification task, reduces the input image of 32×32 into a 120 vector that is called *the feature vector*. Thus,the feature vector can be used with any type of trainable classifier to solve the classification task. In fact, this approach is used in [19], where LeNet-5 is used as a black box feature extractor for several Support Vector Machines (SVM) that are trained based on it. Something similar is performed by [14], where the authors build an AlexNet [13] like model, trained it on a large dataset, and then use the trained model as a feature extractor to train new classifiers; however, their approach is more similar to a transfer learning set-up, as it will be discussed in Section 2.2.

The deepness in the number of layers of the ConvNets has been increasing in order to obtain better results since year 2012. However, the number of layers cannot be increased indefinitely due to the vanishing and exploding gradient problems. In order to overcome this problem, the structure of the traditional ConvNet, comprised basically of convolutional and pooling layers, has been revisited. An example of those architecture designs are found in Residual Networks[18], MobileNet[2], Inception[20] and NasNet[3] ConvNets. Also the study of regularization functions has been of aid in avoiding the overfitting of the Networks in training [21, 22]. For a review of the state of the art in Convolutional Neural Networks the reader may review the works of: [12, 15, 23, 24].

Thus, ConvNets have some advantages compared with traditional artificial neural networks (ANN): reduction of training parameters by shared weights, local connections and object location invariance [12]. An ANN depends on all the connections between its layers; which increases the number of parameters to train, and makes the training of the model more expensive computationally. On the contrary, a ConvNet reduces the number of parameters trained because the convolution operation is of the local type. Yet, the main disadvantages of the ConvNet model are: the training time, which may be large since it is an hyper-parametrized model, and the susceptibility to overfitting. The most important aspect of ConvNets, as a deep learning method, consists in being a *Representation Learning* method that automatically discovers a representation of the data that is used for classification and detection tasks [12], as previously discussed. This is specially important because it implies that the need for carefully hand engineer feature extractors is not required. Remind that the feature extractor is embedded in the design of the ConvNet and, therefore, learned from data.

The success of a ConvNet relies in its architecture as well as on training the model with enough number of samples. Unfortunately, public mammogram datasets do not have as many examples as the number used per category in the ImageNet dataset. In order to overcome this difficulty, TL and FT are studied as a means to train deep ConvNets in order to classify mammogram abnormalities. In this Section we review the concept of TL and FT as it is used in this article.

2.2 Transfer Learning

A definition of the term is found in [25] and [26]. In their work, the purpose of TL is defined as to improve the performance of a learning algorithm in a target learning task $\mathcal{T}^T$ (i.e. pathology classification) over a target domain $\mathcal{D}^T$ (i.e. mammogram ROI images) by using

the knowledge of the learning algorithm trained in a source learning task $\mathcal{T}^S$ (i.e. 1000 category classification) over a source domain $\mathcal{D}^S$ (i.e. natural images) which is larger than the target domain where:

$$\mathcal{D}^S \neq \mathcal{D}^T \quad (1)$$

$$\mathcal{T}^S \neq \mathcal{T}^T \quad (2)$$

Depending on the relations defined by (1) and (2), different categories of TL are defined in literature. However, it is important to consider that computer vision tasks are particular and different from their corresponding data mining tasks. Thus, despite the fact that a mammogram image is very different from a natural image, the visual properties of objects in an image are general (i.e. edges, textures, shapes, etc.).

In a ConvNet, the knowledge is represented by the value of the weights trained by the back-propagation algorithm on each layer. Therefore, TL implies using the pre-trained ConvNet as a feature extractor or replacing the last original layer with a set of layers that are trained to obtain the target learning task desired. The latter is the approach that we have followed in our experiments. One of the main advantages in this technique is that training time is reduced by not re-training the whole ConvNet, but only the added layers or a trainable classifier in the case of using the pre-trained ConvNet as a feature extractor itself.

2.3 Fine Tuning

In this case, some of the last layers of the pre-trained ConvNet are re-trained with the new images $\mathcal{I} \in \mathcal{D}^T$. Thus, the ConvNet is divided in two parts. Let us define γ as the layer from which the re-training of the ConvNet will occur. If the original ConvNet model has L layers, similarly to TL, we can replace the last original layer and add some layers in order to obtain the desired target learning task. Differently to TL, we also choose a r number of layers before the last one that are also to be trained. This means that the original weights from layer 0 to layer $\gamma - 1$ are preserved or frozen. FT presents more computing resources and time training since the number of parameters to be trained is increased by the r layers that are added to the training queue.

2.4 Over and Underfitting

A machine learning algorithm may suffer of two problems when training: overfitting and underfitting. The former reduces the capacity of the model to predict new unseen data which means that the model has a high variance. The latter, means that the model is not complex enough to reflect the nature of the data and find a pattern [27]. ConvNet models are characterized by being over-parameterized; which means that the parameters of the model exceed the size of the training data [28]. The overfitting problem is reflected when plotting the train vs validation accuracy curve of the model. The difference between the train and validation curve should be minimum. In order to overcome overfitting in ConvNets, Data Augmentation, Regularization, and Early Stopping are usually used. Data augmentation is a basic strategy that consists in increasing the size of the dataset by performing transformations to the original images (i.e. rotation, zoom, reflection, etc). On the other hand, regularization techniques aim to penalize extreme parameter weights values (e.g. L_2 regularization) [27] or controlling the co-adaptation between neurons (e.g. Dropout) [13]. Early Stopping is also considered a regularization technique which aims to interrupt training when the performance of the ConvNet degrades on the validation set. This prevents that the model learn a form of statistical noise [29].

As discussed earlier, TL and FT may also prevent overfitting since the CovNet model is not whole retrained; in other words, TL and FT have less parameters to learn compared to training a model from randomly initialized weights. In the present work, Dropout [30] and Early Stopping have been used as regularization techniques altogether with data augmentation; these are changes introduced in this work that differ from our previous experiment.

3 Related Works

3.1 Search Process

In our previous work [1], we used the methodology by [31], in order to find relevant works for study. Table 1 shows the search string designed to retrieve information from: Springer Link, Science Direct (Elsevier), IEEE Xplore, Scopus, Web of Science, ACM digital library, and PubMed. A total of 174 studies (including our previous work) were gathered from each repository as shown in Table 2. From these studies, a total of 32 primary documents were retrieved according to a selection study process where documents should have experimental methodology with results regarding the use of transfer learning in mammogram breast cancer classification.

Table 1: Search String

"breast cancer" **AND** ("classification" **OR** "detection" **OR** "prediction") **AND** ("ensemble learning" **OR** "transfer learning") **AND** mammo*

Table 2: Search Results

Database	**#Publications: 2014-2018**
Scopus	38
IEEE	12
Science Direct	51
PubMed	10
ACM	25
Web of Science	23
Springer	15
Total	**174**

3.2 Literature review discussion

In our literature review, the most common ConvNet used for transfer learning is the model proposed in [13], named *AlexNet*, with a total

of 12 cases. The second most frequent model found is VGG16. These results are shown in Figure 1.

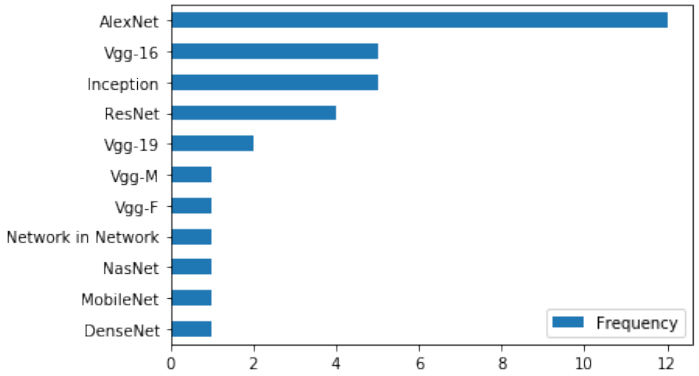


Figure 1: ConvNets used in Literature. This bar plot shows the frequency of the pre-trained ConvNets used in the primary documents retrieved for our literature review. The most used model corresponds to AlexNet (12 cases). At a second place are VGG16 and Inception (5 cases each). Next is ResNet (4 cases) and VGG19 (2 cases).

In mammogram mass abnormality classification and detection there are two main approaches found in literature: a) Processing the whole mammogram image and b) processing the region of interest. The former is found in [32, 33]. Their aim is to find an "*end to end design*". Our approach belongs to the latter, where the ROI image is extracted. In fact processing a whole mammogram images in their original size seems to be a problem itself because mammogram images far exceed the traditional size used in many trained ConvNets in the ImageNet dataset. An interesting approach for an end to end design is proposed in [34]; In their work, the principles of the YOLO [35] architecture are used. However, the mammogram is also resized to 448×448.

As in the case of mammogram classification, Transfer Learning and Fine Tuning are used differently by different authors in literature. In the next subsection this differences are enlightened and discussed.

3.2.1 *Transfer Learning as a Feature Extractor*

In this case, the pre-trained ConvNet is used to extract a feature vector which is later used to train another kind of classifier algorithm like Support Vector Machines (SVM). This case is illustrated in [36]; the author extracts several feature vectors from different layers of the pre-trained AlexNet and trains Support Vector Machines (SVM) for each case. In the end, the author builds an ensemble of SVM. Other similar examples are found in [37, 38].

3.2.2 *Transfer Learning as a new ConvNet Classifier*

In this case, the last full connecting layer of the pre-trained ConvNet may be substituted with a set of additional layers, where the last full connecting layer has only one neuron and the logistic regression for binary classification, or just the number of random initialized neurons required in proportion to the new classification task. For instance, in the case of benign vs malignant classification of the mammogram abnormality this can be achieved with a single neuron or two. Only the added layers are trained while the rest of the ConvNet's weights remain frozen. This approach is tested in both our current and previous work [1]. In a similar fashion, VGG16[39], GoogLeNet[40] and AlexNet[13] are trained in TL in [41].

3.2.3 *Transfer Learning as weight initialization*

In this case the whole ConvNet is re-train but uses the values of the ImageNet pre-trained ConvNet model as initial values for the weights. The last full connecting layer with 1000 categories is substituted by one or two neurons to address the binary classification problem [42]-[43].

3.2.4 *Fine Tuning*

This is the most common technique found in literature. In this case, the model's last full connecting layer is substituted with the number of neurons needed for the new classification or a set of new layers are added before the output layer. Differently to *transfer learning as a new ConvNet*, some of the last layers of the model are re-trained with the new data as indicated in Section 2.3. For example, VGG16[39], InceptionV3[20], and ResNet50 [18] are fine tuned in [44]; the author found that when the number of convolutional blocks exceeds 2, the accuracy of the fine tuned model drops. Also, a comparison of the classification performance between the training of VGG16 in FT and using it as a feature extractor is explored in [45].

3.2.5 *Data Augmentation and Pre-processing*

In literature there is some discussion about the impact of both data augmentation and pre-processing of the medical image. As stated by [46], the achievements in medical images visual tasks with deep learning do not only rely in the ConvNet model but also in the pre-processing of images. For instance, some of the pre-processing methods found in literature are: global contrast normalization (GCN), local contrast normalization, and Otsu's threshold segmentation. However, there is some discussion about improving the image quality. In [37] is reported that global contrast normalization did not aid in improving the experimental results presented in the paper.

Since datasets are not so large, data augmentation is used by almost all researchers. Some of the most common techniques used are: rotations and cropping. However, the rotation operation yields to distortion of the original image. Because of this reason, right angle rotations are preferred to random rotation angles [37, 43, 44].

4 Proposed Approach

4.1 *Transfer Learning and Fine Tuning Model*

The problem to solve is to classify ROI patch mammogram images $\mathcal{I}$ in two classes $\mathcal{Y} = \{benign, malignant\}$. In a supervised learning paradigm this means to find a prediction function $\phi(\cdot)$ that maps an input space $\mathcal{X}$ formed by ROI patch mammogram images ($\mathcal{I} \in \mathcal{X}$) to the output space $\mathcal{Y}$ as indicated in (3)

$$\mathcal{Y} = \phi(\mathcal{X}) \tag{3}$$

However, since TL and FT are to be used to improve the performance of $\phi(\cdot)$, 3 may be written as indicated in 4

$$\mathcal{Y}^T = \phi^{\mathrm{T}}(\mathcal{X}^T) \tag{4}$$

Function $\phi^{\mathrm{T}}(\cdot)$ is to be trained by using pre-trained ConvNets on the ImageNet dataset, which is denominated as $\mathcal{X}^S$. Therefore, our approach satisfies the relations indicated in (1) and (2), since both images and the classification task are different between source and target.

Our approach in both TL and FT consists in replacing the last full connecting layer related to the original ImageNet classification task with a set of layers as indicated in Table 3. The global average pooling (GAvg) layer helps to flatten the original model layer previous to the 1000 full connecting SoftMax classification. The classification layer is comprised of 1 neuron and the Sigmoid function.

Table 3: TL and FT Output

Global Average Pooling 2D
Full Connecting
Dropout
Classification Layer

In TL, only the last layers indicated in Table 3 are to be trained. In FT, we define the γ value that indicates the layer from which the training of the weights is to be performed. It is important to remember that all weights before γ remain with their original value from the ImageNet.

4.2 Dataset

In the present study, as well as our previous work, we use the Curated Breast Imaging Subset of DDSM (CBIS-DDSM)[4] which is an updated and standardized version of the Digital Database for Screening Mammography (DDSM). The dataset includes a subset of the DDSM data selected and curated by a trained mammographer. For our experiments we extract the ROI images from the mammogram images. We have only considered mass problems, leaving micro calcifications for a future work. The dataset is originally organized in train and test sets. The number of images per abnormality class and set type is shown in Table 4.

Table 4: Mass Images in CBIS-DDSM

Dataset Set Type	Benign	Malignant
Train	681	637
Test	231	147

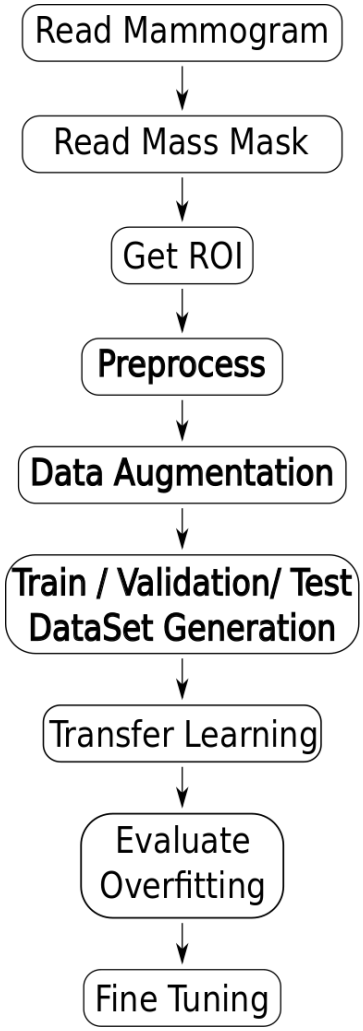


Figure 2: A diagram depicting the methodology followed in our research. First, mammogram and corresponding mass binary mask are read. The mask is used to extract the ROI from the mammogram image. Next, the ROI image is enhanced (pre-process). As a third step, data augmentation is used to increase the number of samples and create train, validation and test sets. Finally, transfer learning is performed and evaluated. The best non overfitting model is selected to be fine tunned

4.3 Methodology

The methodology followed in the current experiments is shown in Figure 2. First, the ROI images are extracted from the mammogram image by using the binary segmentation masks provided in the CBIS-DDSM dataset. After that, images are pre-processed to enhance contrast. Next, a single dataset is formed in order to use data augmentation and create three sets of data: train, validation and test. In this work, we do not use Otsu algorithm to segment the previously obtained ROI by creating an intermediate binary mask. This is because our previous work showed that training with the ROI image segmented with Otsu did not overcome the results obtained with the original background image. Finally, the models are trained in TL and FT, and their performance is evaluated. These steps are described in more detail below.

4.3.1 Image Pre-processing

In this section we present the steps performed in the pre-processing stage of our proposed method. The original mammogram image

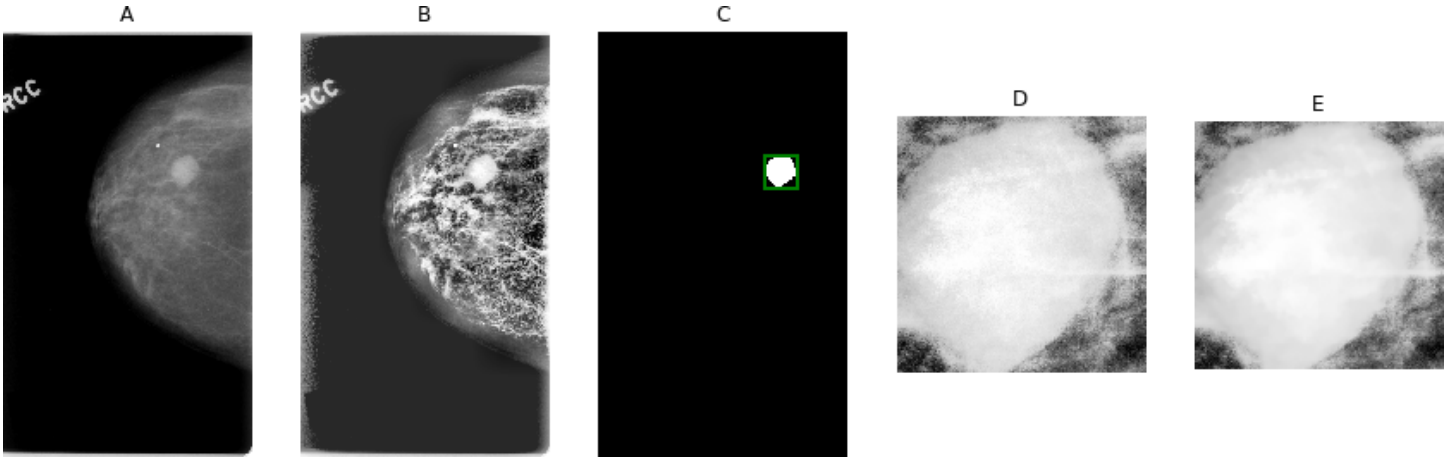


Figure 3: Image Preprocessing. A. The original CBIS-DDSM ROI mammogram image (4704×2744) normalized in UINT 8 resolution ($pixel_value \in [0, 255]$). B. Mammogram image contrast is improved by using The CLAHE equalization algorithm. C. The binary mass mask provided in the dataset is used to find the ROI. D. The ROI is extracted from B step with a size corresponding to the Bounding Box (358×346) found in the binary mask. E. The ROI image in previous step is resized to 328×328, where we used a 8 pixel padding. Finally, the image is cropped around the center to obtain the final desired training dimensions for the ROI image of 320×320 and Fast Non Local Means Denoising Algorithm is used to further filter ROI image scanning noise.

has a pixel value which ranges up to 65 535 (i.e. 16 bit resolution). In order to generate a dataset of PNG images saved in disk, we change the resolution of the images by normalizing them between 0 and 255, considering the minimum and maximum values of pixel in the original DICOM mammogram image. After that, and different to our previous work, we used the Contrast Limited Adaptive Histogram Equalization (CLAHE)[47] to improve image quality. By using the provided binary masks, we extract the ROI through the coordinates of a bounding box around the suspicious mass. A second normalization of the pixel value is carried out on the ROI considering the minimum and maximum pixel values in it. This originates ROI images with different width and height sizes. As in our previous work, aspect ratio is considered. In (5), the aspect ratio is defined; where r is the aspect ratio, w and h are the width and height of the image respectively.

$$r = w/h \tag{5}$$

Aspect ratio was considered previous to resizing the image in order to preserve the best quality possible from the original image in both, upsampling and downsampling procedures. For upsampling, cubic interpolation was used, whereas for downsampling, area interpolation gives best results. Also, images with an aspect ratio inferior to 0.4 and superior to 1.5 where removed from dataset.

$$\mathcal{I} = \{I | I \in \mathbb{R}^{w \times h}, 0.4 \leq r \leq 1.5\} \tag{6}$$

ROI images were resized to a final size of 320×320. This was achieved only with images whose aspect ratio was inside the limits presented in (6). Resizing the ROI images consisted in two parts: 1) we resized the ROI image to 328×328, 2) cropping around the center of the image. In other words, we have resized the original ROI image to a bigger size (with a padding of 8 pixels) and then cropped in the center of the image to obtain the desired size. Finally, the image is filtered with fast non local means denoising algorithm [48]. This is because CBIS-DDSM images are film mammography and appear to have some noise that has not been removed from the image. Figure 3 presents the steps carried out on image pre-processing for a sample of the training set. Both, the full mammogram image and its corresponding binary mass mask are shown in sub-figures A and C respectively. Image C shows the identification of the bounding box. B presents the application of CLAHE on the mammogram image. Finally, E presents the processed ROI image after performing crop center on D which was extracted from B through the bounding box.

4.3.2 Data Augmentation and Dataset Generation

In our previous work we did not use data augmentation and our models presented overfitting. In order to overcome this difficulty, we implemented the output structure indicated in Table 3 which uses dropout. As pointed out in Section 3.2.5, performing transformations over the images distorts them. Because of that, some researchers use right angles.

Table 5: Augmentation Operations

Operation	Probability	Parameters
Rotation	95%	Max Angle 15°
Shear	60%	Max value 25°
Histogram Equalization	60%	
Horizontal Flip	70%	
Bright	80%	min value: 0.6 max value: 1.2
Zoom	100%	min value: 1.0 max value: 1.2

In our case, we have used the Augmentor Library [49], which has been designed to permit rotations of the images limiting the degree of distortion. Additionally, the Augmentor library permits to apply other operations for data augmentation like zoom, bright, shear. The function uses a probability value to control the number of artificially created images. To augment the dataset, we first join both Train and Test sets. The dataset was increased to a total of 60 000 images, where 80% of the images are used for training, 10%

for validation and 10% for testing. The augmentation operations used are depicted in Table 5.

4.3.3 Model Training

The generated augmented dataset is tested first in Transfer Learning. A total of 4096 neurons are used for the FC layer, a dropout value of 0.2, and a single neuron in the output layer (or classification layer, see Table 3) for binary classification. Early Stopping, with a patience of 50, was enabled in order to stop training when the performance degrades. A maximum number of 1 000 epochs is proposed. The learning rate for TL is 1×10^{-5}. The loss function is set to binary cross entropy and the optimization algorithm used is RMSProp [50]. Binary cross entropy was chosen because the classification layer of our model uses the Sigmoid activation function to discriminate between benign and malignant mass pathology by using one neuron. RMSProp, an adaptive gradient algorithm, is frequently used in computer vision tasks [51]. For instance, the results achieved in [52] indicate a better training result obtained by using RMSProp instead of Stochastic Gradient Descent (SGD). Also, as presented in [53], RMSProp outperforms other common optimization algorithms [51]. This has inspired the theoretical research in [51], where the authors establish the reasons of the success of RMSProp in deep learning and propose new algorithms for optimization. In our study, we consider that RMSProp is convenient for our image classification task because of the aforementioned reasons.

Transfer Learning was carried out in 20 models provided by the Keras API in Tensorflow v1.13.1 [54]. The models used are shown in Table 6.

Table 6: Pre trained ConvNets used

Vgg	**Densenet**	**Inception**	**Nasnet**
16, 19	121, 169, 201	v3, Resnet-v2	large, mobile
Mobilenet	**Resnet**	**Xception**	**Resnext**
v1, v2	50, 50v2, 101, 101v2, 152, 152v2	v1	50, 101

Each model performance is evaluated in the train-validation curve of accuracy. If the difference between train accuracy and test accuracy is over 10%, we consider that overfitting has occurred and the model is rejected. Once a suitable not overfitting model is found, FT is performed in order to further increase the performance of the selected classifier.

FT training parameters are: learning rate of 2×10^{-7}; Dropout remains at 0.2. Respect to the FC neurons, only Global average pooling is performed. Binary cross entropy is set as the cost function and the optimization algorithm used is also RMSProp.

4.3.4 Computing Resources

In order to train deep learning models it is necessary to use Graphical Processing Units (GPU). In our case we have used: Nvidia Tesla K80, with 12 GB of memory, Nvidia Tesla K40, with 12 GB of memory, and Nvidia GeForce RTX 2080, with 8 GB of memory.

Regarding software resources, our program used Tensorflow v13.1 [54] as the machine learning framework. For image data augmentation, as indicated before, the Augmentor library [49] was used.

5 Experimental Results

In this Section, we present our experimental results. As a difference wrt. to our previous work, we include additional metrics to evaluate the performance of the classification model such as the area under the ROC curve and the F_1 Score. These metrics are described below. According to the methodology proposed in Figure 2, it is important to estimate the overfitting of the trained model $\phi^{\mathrm{T}}(\cdot)$. In order to do so, let us define β as the overfitting ratio by comparing train (*train_acc*) and validation accuracy (*valid_acc*) as in

$$\beta = \frac{train_acc}{valid_acc} \tag{7}$$

If $\beta \approx 1$, we could say that *train_acc* $\approx$ *valid_acc* and therefore that there is little overfitting. Values of $\beta > 1$ will reflect that there is a considerable difference between train and validation accuracy, meaning that the model has overfitted.

5.1 Performance Metrics

The confusion matrix compares both the prediction of the trained classifier and the true labels provided in the test set. It consists of four main measures: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The total positive cases are $P = TP + FN$; similarly, the total negative cases: $N = TN + FP$. From these measures, the true positive rate (8), false positive rate (9), and true negative rate (10) are derived.

$$TPR = \frac{TP}{P} \tag{8}$$

$$FPR = \frac{FP}{N} \tag{9}$$

$$TNR = \frac{TN}{N} \tag{10}$$

The elements over the diagonal belong to TP and TN and reflect all the correct classifications made by the model. The FP and FN correspond to wrongly classified cases. For instance, FN corresponds to a true malignant tumor that is classified as benign. From these measurements, metrics like: Accuracy (11), F_1 Score (12), and the area under the ROC curve (13) are defined.

$$ACC = \frac{TP + TN}{TP + TN + FP + TN} \tag{11}$$

$$F_1 = \frac{2TP}{2TP + FP + FN} \tag{12}$$

$$AUC = \frac{1}{2}\left(1 + TPR - FPR\right) \tag{13}$$

5.2 Transfer Learning Experiment

The ImageNet pre-trained ConvNets presented in Table 6 are trained under TL to predict mammogram abnormalities classes in mammogram roi images.

Table 7: Transfer Learning Results

Model	*AUC*	F_1Score	*ACC*
resnet-50-TL	0.861	0.858	0.861
mobilenet-TL	0.854	0.855	0.854
resnet-152-TL	0.853	0.847	0.853
resnet-101-TL	0.846	0.848	0.846
resnet-152v2-TL	0.844	0.845	0.844
resnet-50v2-TL	0.843	0.854	0.843
resnet-101v2-TL	0.832	0.833	0.832
densenet-201-TL	0.821	0.823	0.821
xception-TL	0.816	0.82	0.816
densenet-169-TL	0.812	0.816	0.812
resnext-101-TL	0.806	0.82	0.806
nasnet-l-TL	0.804	0.802	0.804
mobilenet-v2-TL	0.791	0.801	0.791
densenet-121-TL	0.79	0.79	0.789
inception-v3-TL	0.778	0.781	0.779
resnext-50-TL	0.777	0.764	0.777
inception-resnet-v2-TL	0.744	0.746	0.744
nasnet-m-TL	0.712	0.73	0.712
vgg16-TL	0.644	0.654	0.644
vgg19-TL	0.629	0.659	0.629

Table 8: Transfer Learning Overfitting Ratio β

Model	*train_acc*	*test_acc*	β
vgg16-TL	0.65	0.64	1.01
vgg19-TL	0.64	0.63	1.02
resnet-50-TL	1.00	0.86	1.16
mobilenet-TL	1.00	0.85	1.17
resnet-152-TL	1.00	0.85	1.17
resnet-101-TL	1.00	0.85	1.18
resnet-152v2-TL	1.00	0.84	1.19
resnet-50v2-TL	1.00	0.84	1.19
resnet-101v2-TL	1.00	0.83	1.20
inception-resnet-v2-TL	0.90	0.74	1.21
densenet-201-TL	1.00	0.82	1.22
xception-TL	1.00	0.82	1.22
resnext-101-TL	0.99	0.81	1.22
densenet-169-TL	1.00	0.81	1.23
densenet-121-TL	0.98	0.79	1.24
nasnet-l-TL	1.00	0.80	1.24
mobilenet-v2-TL	1.00	0.79	1.26
resnext-50-TL	0.98	0.78	1.26
inception-v3-TL	0.99	0.78	1.27
nasnet-m-TL	0.91	0.71	1.28

As indicated in Section 4.3.3, 4096 output neurons and a dropout rate of 0.2 are used. The results achieved are shown in Table 7. The overfitting ratio stated in (7) is presented in Table 8. The results indicate that Resnet-50 has the best AUC, however its β shows that there is overfitting. On the contrary, both VGG16 and VGG19 do not present overfitting, but their classification performance is lower compared to Resnet-50. These results are reflected in Figure 4 and 5, where the plot train vs test accuracy is presented for Resnet-50 and VGG16, respectively.

5.3 Fine Tuning Experiment

According to the results indicated in Tables 7 and 8, we proceeded to train the VGG16 in Fine Tuning. Different deepness levels where tried in order to search for classification performance improvement. This is indicated through the γ value. VGG16 was trained from layers $\gamma = 8$, $\gamma = 10$. In order to denominate the trained model, we propose to use the pre-trained ConvNet name followd by the layer from which fine tunning occured and added the keyword FT to distinguish the model from those trained in TL mode. For instance, *VGG16-10-FT* means that VGG16 was fine tuned from layer 10. Similarly, we trained VGG19 at $\gamma = 17$. In all cases, only global average pooling followed by dropout and the classification layer were used; except for the case of *VGG16-8-FT*, where a full connecting layer of 4096 neurons was used.

The results obtained are shown in Table 9. We have complimented the experimental results with the Fine Tuning of models: Xception, Resnet101, Resnet152 and Resnet50. It is observable that the best results are achieved by VGG models. The best result achieved corresponds to the VGG16-8-FT. However, Table 10 suggest that the second best result (VGG16-10-FT) has less overfitting and therefore is preferred. In fact β value for VGG16-10-FT and VGG19-17-FT is similar, but the performance of the latter is poorer.

In Figure 6, the plot of train and validation accuracy vs the number of epochs for VGG16-10-FT is presented. Figure 7 shows the ROC curve obtained, whereas Figure 8 presents the confusion matrix for the test set generated.

Table 9: Fine Tuning Results

Model	*AUC*	F_1Score	*ACC*
vgg16-8-FT	0.844	0.85	0.844
vgg16-10-FT	0.816	0.822	0.816
vgg19-17-FT	0.774	0.774	0.774
xception-127-FT	0.571	0.664	0.571
xception-122-FT	0.526	0.67	0.526
xception-130-FT	0.504	0.663	0.504
resnet-101-343-FT	0.5	0.667	0.5
resnet-101-340-FT	0.5	0.667	0.5
resnet-152-510-FT	0.5	0.667	0.5
resnet-50-173-FT	0.5	0.667	0.5
resnet-101-330-FT	0.5	0	0.5
resnet-152-513-FT	0.5	0	0.5
resnet-50-160-FT	0.5	0	0.5
resnet-50-170-FT	0.5	0	0.5

Table 10: Fine Tuning Overfitting Ratio β

Model	*train_acc*	*test_acc*	β
vgg19-17-FT	0.82	0.77	1.07
vgg16-10-FT	0.87	0.82	1.07
vgg16-8-FT	0.92	0.84	1.09

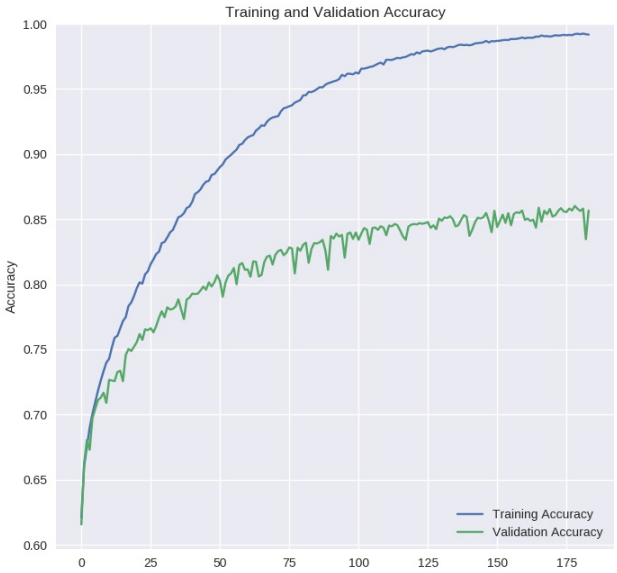


Figure 4: Comparison of the train and validation accuracy for **Resnet50-TL** when trained in transfer learning. As can be seen, despite that Resnet50 achieves an $ACC = 0.86$ accuracy, the model presents overfitting due to the distance between both curves, described by the ratio $\beta = 1.16$

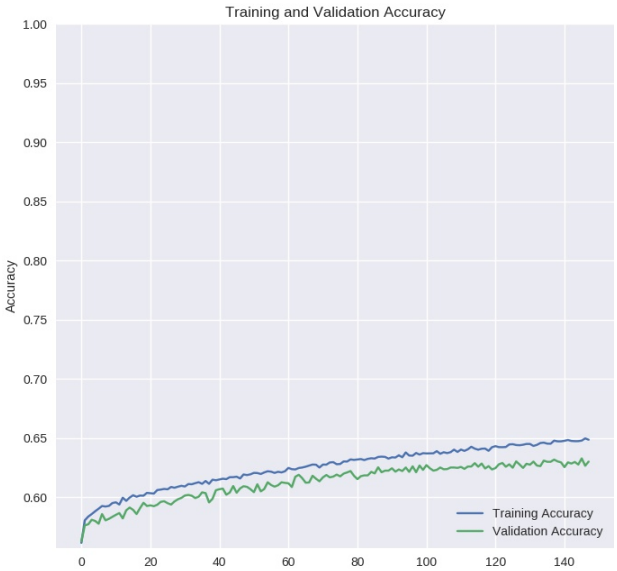


Figure 5: Comparison of the train and validation accuracy for **Vgg16-TL** when trained in transfer learning. VGG16 achieves a lower level of accuracy compared to Resnet50 ($ACC = 0.64, \beta = 1.01$) but there is no overfitting, due to the little distance between train and validation curves.



Figure 6: Fine Tuning VGG16 achieves the best result for our dataset based on CBIS-DDSM ($ACC = 0.87, \beta = 1.07$). Fine Tuning has overcome overfitting and managed to increase classification performance as comparing the train and validation curves for the accuracy metric show in this figure.

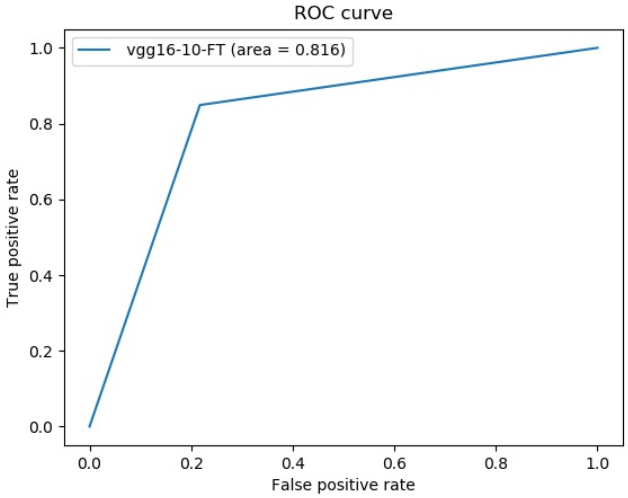


Figure 7: ROC Curve for the binary classification of ROI mammogram images test set. It depicts the performance of the Fine Tuned VGG16-10-FT model, which achieved $AUC = 0.818$

Fine Tuning the VGG16 model from layer 10 ($\gamma = 10$) helps to increase the performance of the classifier while controlling the overfitting ($\beta = 1.07$), which means that train accuracy is 7% over the test accuracy.

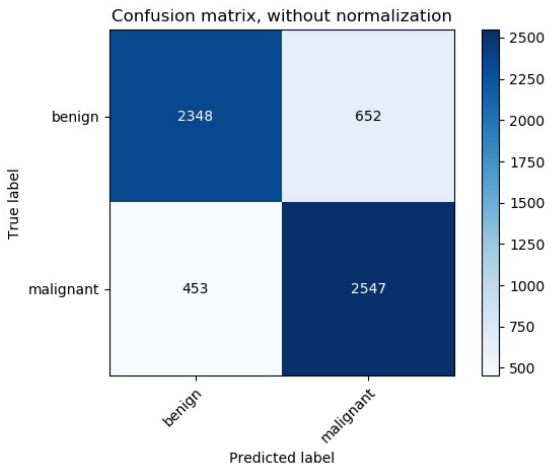


Figure 8: Confusion matrix of the VGG16-10-FT over the test set with 3000 images per category (malignant, benign). In this matrix, True Positive cases are 2547 (84.9%) and True Negatives are 2348(78.2%). False Positive are 652 (21.7%) and False Negatives are 453 (15.1%). The VGG16 fine tuned model has been able to reduce false negative cases.

6 Conclusion

The present study compared the performance of TL and FT of different pre-trained ConvNet models on the ImageNet dataset such as: VGG, DenseNet, Inception, Resnet, Resnext and Xception. In our previous work, we have experimented with Nasnet and Mobilenet without data augmentation. This showed that the models had a trend to overfit. In order to overcome this problem and increase the performance of a Fine Tuned model, in this work we have used data augmentation as indicated in Section 4.3.2. Our experiments showed that increasing the dataset up to 30 000 images per category helped to achieve good results. Special care was taken to increase the dataset by using the Augmentor library [49] which permits to rotate images avoiding excessive distortion. Compared to our previous work, the image pre-processing has also been changed: instead of using histogram equalization, CLAHE was selected to enhance the contrast the images. Also, the image was resized to a bigger value in order to crop around the center. Finally, image filtering was applied to reduce the presence of noise in the image, as indicated in Section 4.3.1.

In order to estimate overfitting, we proposed a simple ratio relation described in (7). This permitted to conclude that models could achieve good results in classification metrics but overfit in the end. Considering this, we decided to increase the complexity of the model by using the fine tuning technique, where weights from layer 1 to $\gamma - 1$ are frozen, and weights from layer γ until the end are trained with back-propagation. This allowed to increase the performance of both VGG16 and VGG19. Increasing the number of neurons in the FC layer of VGG16-8-FT improved the results but there is a slight overfitting.

The CBIS-DDSM dataset, despite the fact of having at most 1 696 examples for mass abnormality classification, their samples present some artifacts and noise, which is the reason why we used some pre-processing and filtering algorithms in order to improve the image. This is probably due to the fact that the original images are of film type. Certainly, our experiments suggest that TL and FT of pre-trained ConvNets is able to classify film mammogram ROI images. However, the data augmentation increase of the original dataset is considerably. Therefore, public mammogram datasets of Full Digital Mammography, which has better quality than film, with enough sample data would aid to train better classifiers.

With respect to our previous work, we have also included new metrics to evaluate the performance of the classifier. This is important, since the accuracy metric is susceptible to be distorted when the dataset is skewed or unbalance. In other words, it tends to benefit the class with a majority of examples.

7 Future Works

The interest in improving CAD systems in mammography is clear since the disease is a public health problem with high rates of both incidence and mortality. In this and our previous work, we have used the CBIS-DDSM dataset mainly. In a future work, we aim to evaluate the performance of TL and FT in other datasets such as INbreast[55] and Mias[56].

Also, we will be addressing the problem of localization and detection of the mass in the mammogram image. This problem is also of interest and could be formulated from the classification problem here addressed. One of the important things to notice is the peculiarity of the size of the mammogram image compared to the size used in ImageNet trained ConvNets. Mammogram images are of considerable size and it could be of interest to design both the classifier and the object detection avoiding to excessively resize the original image.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment This research was carried out using the research computing facilities offered by Scientific Computing Laboratory of the Research Center on Mathematical Modeling: MODEMAT, Escuela Politécnica Nacional - Quito and those also provided by Corporación Ecuatoriana para el Desarrollo de la Investigación y la Academia (CEDIA). The authors also gratefully acknowledge the financial support provided by the Escuela Politécnica Nacional, for the development of the project PREDU 2016-013.

References

[1] L. G. Falconí, M. Pérez, and W. G. Aguilar. Transfer learning in breast mammogram abnormalities classification with mobilenet and nasnet. In *2019 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 109–114, June 2019. doi: 10.1109/IWSSIP.2019.8787295.

[2] Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017.

[3] Barret Zoph, Vijay Vasudevan, Jonathon Shlens, and Quoc V Le. Learning transferable architectures for scalable image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8697–8710, 2018.

[4] Rebecca Sawyer Lee, Francisco Gimenez, Assaf Hoogi, Kanae Kawai Miyake, Mia Gorovoy, and Daniel L. Rubin. A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific Data*, 4:170177 EP –, Dec 2017. URL https://doi.org/10.1038/sdata.2017.177. Data Descriptor.

[5] Kiven Erique Lukong. Understanding breast cancer–the long and winding road. *BBA clinical*, 7:64–77, 2017.

[6] World health organization. global health observatory. geneva: World health organization; 2018. who.int/gho/database/en/. 2018.

[7] Jennifer S Drukteinis, Blaise P Mooney, Chris I Flowers, and Robert A Gatenby. Beyond mammography: new frontiers in breast cancer screening. *The American journal of medicine*, 126(6):472–479, 2013.

[8] D. Selvathi and A. Aarthypoornila. Performance analysis of various classifiers on deep learning network for breast cancer detection. volume 2018-Janua, pages 359–363. ISBN 9781509067305. doi: 10.1109/CSPC.2017.8305869.

[9] Afsaneh Jalalian, Syamsiah Mashohor, Rozi Mahmud, Babak Karasfi, Abdul Rahman B Ramli, Communication Systems Engineering, Universiti Putra, Health Science, Universiti Putra, and Qazvin Branch. Review article : FOUNDATION AND METHODOLOGIES IN COMPUTER-AIDED. pages 113–137, 2017.

[10] X. Zhang, Y. Zhang, E.Y. Y Han, N. Jacobs, Q. Han, X. Wang, and J. Liu. Whole mammogram image classification with convolutional neural networks. volume 2017-Janua, pages 700–704, 2017. ISBN 9781509030491. doi: 10.1109/BIBM.2017.8217738.

[11] John Stoitsis, Ioannis Valavanis, Stavroula G Mougiakakou, Spyretta Golemati, Alexandra Nikita, and Konstantina S Nikita. Computer aided diagnosis based on medical image processing and artificial intelligence methods. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 569(2):591–595, 2006.

[12] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436, 2015.

[13] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, jun 2017. ISSN 15577317. doi: 10.1145/3065386.

[14] L. Hertel, E. Barth, T. Käster, and T. Martinetz. Deep convolutional neural networks as generic feature extractors. In *2015 International Joint Conference on Neural Networks (IJCNN)*, pages 1–4, July 2015. doi: 10.1109/IJCNN.2015.7280683.

[15] Yann LeCun, Yoshua Bengio, et al. Convolutional networks for images, speech, and time series, the handbook of brain theory and neural networks, 1998.

[16] Y Bengio. Convolutional Networks for Images, Speech, and Time-Series Unsupervised Learning of Speech Representations View project Parsing View project. Technical report, 1997. URL https://www.researchgate.net/publication/2453996.

[17] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009. doi: 10.1109/cvprw.2009.5206848.

[18] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. Technical report. URL http://image-net.org/challenges/LSVRC/2015/.

[19] Fabien Lauer, Ching Y. Suen, and Gérard Bloch. A trainable feature extractor for handwritten digit recognition. *Pattern Recognition*, 40(6):1816 – 1824, 2007. ISSN 0031-3203. doi: https://doi.org/10.1016/j.patcog.2006.10.011. URL http://www.sciencedirect.com/science/article/pii/S0031320306004250.

[20] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, and Jonathon Shlens. Rethinking the Inception Architecture for Computer Vision. Technical report. URL https://arxiv.org/pdf/1512.00567.pdf.

[21] Evgeny A. Smirnov, Denis M. Timoshenko, and Serge N. Andrianov. Comparison of Regularization Methods for ImageNet Classification with Deep Convolutional Neural Networks. *AASRI Procedia*, 6:89–94, 2014. ISSN 22126716. doi: 10.1016/j.aasri.2014.05.013.

[22] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, and Ruslan Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Technical report, 2014.

[23] Waseem Rawat and Zenghui Wang. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Computation*, 29(9):2352–2449, sep 2017. ISSN 0899-7667. doi: 10.1162/neco_a_00990. URL http://www.mitpressjournals.org/doi/abs/10.1162/neco{_}a{_}00990.

[24] Neena Aloysius and M. Geetha. A review on deep convolutional neural networks. In *2017 International Conference on Communication and Signal Processing (ICCSP)*, pages 0588–0592. IEEE, apr 2017. ISBN 978-1-5090-3800-8. doi: 10.1109/ICCSP.2017.8286426. URL http://ieeexplore.ieee.org/document/8286426/.

[25] Karl Weiss, Taghi M Khoshgoftaar, and DingDing Wang. A survey of transfer learning. *Journal of Big Data*, 3(1):9, 2016.

[26] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2010.

[27] Patrick Hill and Uma Kanagaratnam. *Python Machine Learning Sebastian Rashka*, volume 58. OUP, 2016.

[28] Mingchen Li, Mahdi Soltanolkotabi, and Samet Oymak. Gradient descent with early stopping is provably robust to label noise for overparameterized neural networks. *CoRR*, abs/1903.11680, 2019. URL http://arxiv.org/abs/1903.11680.

[29] Mingchen Li, Mahdi Soltanolkotabi, and Samet Oymak. Gradient descent with early stopping is provably robust to label noise for overparameterized neural networks. *CoRR*, abs/1903.11680, 2019. URL http://arxiv.org/abs/1903.11680.

[30] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014. URL http://jmlr.org/papers/v15/srivastava14a.html.

[31] Barbara Kitchenham and S Charters. Guidelines for performing Systematic Literature Reviews in Software Engineering. In *Engineering*, volume 2, page 1051. 2007. ISBN 1595933751. doi: 10.1145/1134285.1134500.

[32] Gustavo Carneiro, Jacinto Nascimento, and Andrew P. Bradley. *Deep Learning Models for Classifying Mammogram Exams Containing Unregistered Multi-View Images and Segmentation Maps of Lesions*. Elsevier Inc., 1 edition, 2017. ISBN 9780128104095. doi: 10.1016/B978-0-12-810408-8.00019-5. URL http://dx.doi.org/10.1016/B978-0-12-810408-8.00019-5.

[33] Stephen Morrell, Zbigniew Wojna, Can Son Khoo, Sebastien Ourselin, and Juan Eugenio Iglesias. Large-scale mammography CAD with deformable conv-nets. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 11040 LNCS, pages 64–72, 2018. ISBN 9783030009458. doi: 10.1007/978-3-030-00946-5_7. URL http://link.springer.com/10.1007/978-3-030-00946-5{_}7.

[34] M.A. Al-masni, M.A. Al-antari, J.-M. Park, G. Gi, T.-S. T.-Y. Kim, P. Rivera, E. Valarezo, M.-T. Choi, S.-M. Han, and T.-S. T.-Y. Kim. Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system. *Computer Methods and Programs in Biomedicine*, 157:85–94, 2018. doi: 10.1016/j.cmpb.2018.01.017.

[35] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[36] Benjamin Q Huynh, Hui Li, and Maryellen L Giger. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. *J. Med. Imag*, 3(3):34501, 2016. doi: 10.1117/1.JMI.3.3.034501. URL http://benhuynh.github.io/tl{_}jmi.pdf.

[37] Ana Perre, Luís A. Alexandre, and Luís C. Freire. Lesion classification in mammograms using convolutional neural networks and transfer learning. *Lecture Notes in Computational Vision and Biomechanics*, 27:360–368, jul 2018. ISSN 22129413. doi: 10.1007/978-3-319-68195-5_40. URL https://www.tandfonline.com/doi/full/10.1080/21681163.2018.1498392.

[38] Y. Hu, J. Li, and Z. Jiao. Mammographic mass detection based on saliency with deep features. In *ACM International Conference Proceeding Series*, volume 19-21-Augu, pages 292–297, 2016. ISBN 9781450348508. doi: 10.1145/3007669.3007714.

[39] Karen Simonyan and Andrew Zisserman. VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION. Technical report, 2015. URL http://www.robots.ox.ac.uk/.

[40] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[41] Azam Hamidinekoo, Zobia Suhail, Erika Denton, and Reyer Zwiggelaar. Comparing the performance of various deep networks for binary classification of breast tumours. In *14th International Workshop on Breast Imaging (IWBI 2018)*, page 39, 2018. ISBN 9781510620070. doi: 10.1117/12.2318084.

[42] Oliver Diaz, Robert Marti, Xavier Llado, and Richa Agarwal. Mass detection in mammograms using pre-trained deep learning models. In Elizabeth A. Krupinski, editor, *14th International Workshop on Breast Imaging (IWBI 2018)*, volume 10718, page 12. SPIE, jul 2018. ISBN 9781510620070. doi: 10.1117/12.2317681.

[43] F. Jiang, H. Liu, S. Yu, and Y. Xie. Breast mass lesion classification in mammograms by transfer learning. In *ACM International Conference Proceeding Series*, pages 59–62, 2017. ISBN 9781450348270. doi: 10.1145/3035012.3035022.

[44] Hiba Chougrad, Hamid Zouaki, and Omar Alheyane. Deep Convolutional Neural Networks for breast cancer screening. *Computer Methods and Programs in Biomedicine*, 157:19–30, apr 2018. ISSN 18727565. doi: 10.1016/j.cmpb.2018.01.011. URL https://linkinghub.elsevier.com/retrieve/pii/S0169260717301451.

[45] Shuyue Guan and Murray Loew. Breast Cancer Detection Using Transfer Learning in Convolutional Neural Networks. In *2017 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pages 1–8. IEEE, oct 2017. ISBN 978-1-5386-1235-4. doi: 10.1109/AIPR.2017.8457948. URL https://ieeexplore.ieee.org/document/8457948/.

[46] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. A survey on deep learning in medical image analysis, dec 2017. ISSN 13618423. URL https://linkinghub.elsevier.com/retrieve/pii/S1361841517301135.

[47] Karel Zuiderveld. Contrast limited adaptive histogram equalization. In *Graphics gems IV*, pages 474–485. Academic Press Professional, Inc., 1994.

[48] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. Non-Local Means Denoising. *Image Processing On Line*, 1:208–212, 2011. doi: 10.5201/ipol.2011.bcm_nlm.

[49] Marcus D Bloice, Peter M Roth, and Andreas Holzinger. Biomedical image augmentation using Augmentor. *Bioinformatics*, 04 2019. ISSN 1367-4803. doi: 10.1093/bioinformatics/btz259. URL https://doi.org/10.1093/bioinformatics/btz259.

[50] Geoffrey Hinton, Nitish Srivastava, and Kevin Swersky. Neural networks for machine learning lecture 6a overview of mini-batch gradient descent. *Cited on*, 14(8), 2012.

[51] Mahesh Chandra Mukkamala and Matthias Hein. Variants of rmsprop and adagrad with logarithmic regret bounds. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, ICML'17, page 2545–2553. JMLR.org, 2017.

[52] Andrej Karpathy and Li Fei-Fei. Deep visual-semantic alignments for generating image descriptions. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[53] Tom Schaul, Ioannis Antonoglou, and David Silver. Unit tests for stochastic optimization. *arXiv preprint arXiv:1312.6055*, 2013.

[54] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL http://tensorflow.org/. Software available from tensorflow.org.

[55] Inês C Moreira, Igor Amaral, Inês Domingues, António Cardoso, Maria Joao Cardoso, and Jaime S Cardoso. Inbreast: toward a full-field digital mammographic database. *Academic radiology*, 19(2):236–248, 2012.

[56] P SUCKLING J. The mammographic image analysis society digital mammogram database. *Digital Mammo*, pages 375–386, 1994.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 158-165 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Analysis and Improvement of an Innovative Solution Through Risk Reduction: Application to Home Care for the Elderly

Linda Acosta-Salgado*, Auguste Rakotondranaivo, Eric Bonjour

Research Team on Innovation Process (Laboratoire ERPI, EA N° 3767), Université de Lorraine, Nancy, France

ARTICLE INFO

Article history:
Received: 17 January, 2020
Accepted: 27 March, 2020
Online: 08 May, 2020

Keywords:
Acceptability
Acceptance
Need
Risk
Elderly care
Healthcare innovation

ABSTRACT

The increase in the number of elderly people requires multiple efforts to maintain their well-being and health. A wide variety of products and services have been created to enable seniors to live at home for as long as possible. The market success of these solutions depends on acceptance by the different stakeholders. Older people are reluctant to change their environment, most notably their home. Solutions must provide a benefit and/or reduce risks related to their everyday life in order to be accepted. Design methods are mainly focused on needs analysis, while acceptability assessment models are based on the study of benefits. A need may also correspond to a risk that may be present in the initial situation and that has to be eliminated or reduced. The notion of risk is not sufficiently integrated in these models. This paper proposes a new approach to analyze the actual situation of elderly people at home based on risk analysis. The objective is to contribute to the design of solutions that will be more readily accepted by this population. A model for estimating the risk of falling has been proposed. The probability of two elderly people falling in their home is assessed using this model. The design and improvement of solutions are explored using the results obtained.

1. Introduction

The ageing of the world's population is a key issue for the years to come. Demographic trends reveal an increasing proportion of older people and rising life expectancy. According to projections made by the National Institute of Statistics and Economic Studies (INSEE), in 2070, France should have 76.5 million inhabitants, 10.7 million more than in 2013. Specifically, the population aged 65 and over should increase by 11.2 million and those aged 20–64 decrease by 7.9 million [1]. Thus, by 2070, there will be more seniors than younger people to help them. This situation confronts the health field with a major societal challenge: delaying the age of entry into dependency and improving the well-being of the elderly by controlling risks and expenditure [2].

Public policy encourages the creation of new products and services for keeping people at home as long as possible. For the private industrial and commercial sectors, ageing represents an economic opportunity. Technology companies see this segment of the population as an attractive market for products and services that are easy to use and affordable [3]. In recent years, many technological and organizational innovations have been developed to improve elderly home care. These include video-vigilance systems, fall detectors, tele-assistance and home automation equipment.

Innovative solutions for the ageing population could increase elderly people's capabilities and help them to reduce their entry into dependency by transforming their living environment [4]. Hence, they have to satisfy the needs and preferences of multiple stakeholders (older people, their families, caregivers, care structures, etc.). However, these stakeholders are not always in favor of integrating changes into their daily lives [5].

The success of an innovative project is evaluated by its acceptability in the design phase and by its acceptance in the use phase [6]. To promote the acceptance of a product or service, it is necessary to accompany its design process, from the early stages, by evaluating its acceptability [7]. To improve innovation acceptance, the project must provide benefits and/or reduce the risks perceived by its future users.

In order to improve the well-being of older people at home and increase the likelihood that innovative solutions will be accepted by their various users, it is important to integrate the notion of risk

*Corresponding Author: Linda Acosta-Salgado, Université de Lorraine, Nancy, France. Email: linda.acosta-salgado@univ-lorraine.fr

into the design and improvement process of solutions. The recognition of the positive and negative points of a situation leads design teams to imagine new products and develop new ways of proceeding [8]. Within this framework, this article proposes an analysis approach for improving innovative solutions based on risk analysis of the real situation of elderly people staying at home.

This paper is an extension of work presented in the International Conference on Engineering, Technology and Innovation (ICE/ITMC) [7]. The approach proposed for analyzing and improving a solution based on its risk study is more detailed. After this introduction, the second section presents a literature review on design methods, acceptability assessment models and risks for older people living at home. The next section proposes a conceptual approach and a model to evaluate the risk of falling. It also describes the methods and materials used to assess the fall risk of two people living in their own home. The fourth section presents the results and explores the value of addressing these results to improve a solution. The final part presents the discussion, conclusions and research perspectives.

2. Literature Review

2.1. Design Methods

The notion and discipline of design has been in constant evolution [9]. The development of new products was initially based on the search for function-based solutions. Over time, user needs were seen as a determining factor in the design process, and several user-centered design currents have developed in recent decades.

Today, design approaches such as user innovation, user experience, emotional design and design thinking, seek to better understand the needs of users and conceive new solutions alongside them [10]. Different design stages have been proposed as well as different ways of identifying them [11]. Generally, these processes include phases for understanding the problem, defining the design project, finding solutions, prototyping and evaluation. At the end of this process, the expected results boost the acceptability of the new solutions.

In various design models, the notion of need is fundamental. Need can be defined as a desire resulting from a lack or a dissatisfaction in the goal achievement process [12]. It is relative to an individual and is perceived differently by each person. Need perception can also be the product of risk perception. This need consists in eliminating or reducing the severity or the probability of occurrence of this risk.

2.2. Acceptability Assessment Models

A lot of research has been carried out to assess the acceptability and acceptance of innovative solutions. Researchers have proposed different theoretical models such as the Theory of Reasoned Action (TRA) [13], the Theory of Planned Behavior (TPB) [14] and the Technology Acceptance Model (TAM) [15]. These models have been modified several times by incorporating new factors and relationships between them [16].

The changes made to these models have been aimed at adapting them to evaluate specific innovations in a given context (actors, uses) and in a particular field. Many models have been created in this way, such as TAM 2, TAM 3 and UTAUT.

A literature review of models used in the health field [7] revealed that most existing models are not suitable to assess acceptability at the design stage or to incorporate the preferences of multiple stakeholders. In these models, the assessment is mainly based on the analysis of factors such as utility and ease of use. The review identified only a few studies that incorporate "perceived risk" as a factor in acceptance assessment models [17–19]. No documents were identified proposing a risk-based approach to improve an innovative solution or to design new concepts.

2.3. Risks for Elderly People at Home

Risks can be defined as people's feelings about the potential negative consequences of a situation [20] that may prevent goal achievement. Risk perception is a decisive variable in the choice of behavior [21]. Understanding and managing risk is linked to innovation. Awareness of risks and their factors is a necessary condition for making intelligent decisions in innovation [8]. They are both a threat to the success of a project and an opportunity to make improvements.

Different approaches have been used to diagnose the risks faced by older people at home [22–24]. A literature review conducted by Stuck presents a list of risks of declining functional status for older adults living at home [25]. However, these risks focus on the biological aspects of elderly people such as physical and mental health and social aspects. Other studies have identified additional risks that take into account the impact of the environment on the well-being of elderly people living at home [26–28].

In a previous study [7], risk identification regarding elderly people at home was carried out. This was done through a literature review and the application of expert interviews. Twenty-two risks were identified and classified into eight categories (figure 1).

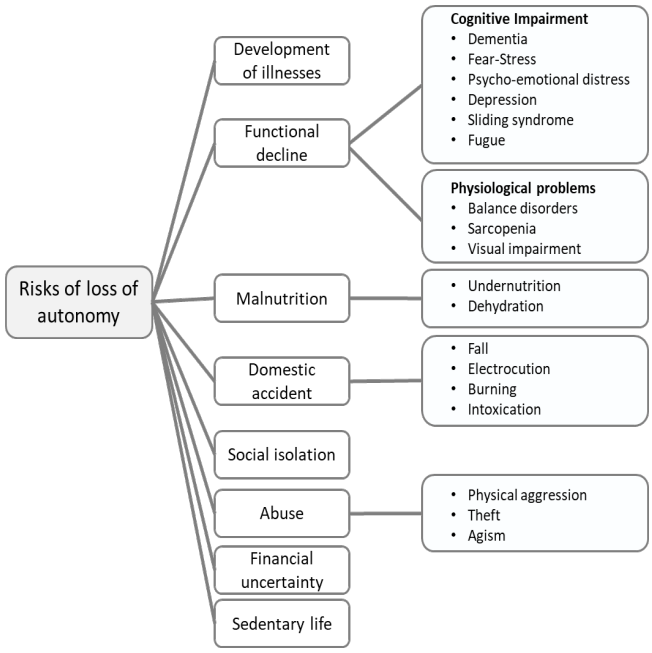


Figure 1: Risks to the elderly at home

The occurrence of such risks is a determining factor in the deterioration of elderly people's well-being and endangers their ability to stay at home. If one of these risks is detected, it could influence the occurrence of others in a systematic way. In order to estimate the likelihood of a risk occurring, it is necessary to identify the causal factors.

2.4. Risk of Falling

Falls are among the most common and serious problems, the consequences of which mainly affect individuals (functional decline, morbidity, institutionalization, mortality, etc.) and the health system [29]. Most falls by the elderly are the result of interactions between intrinsic (physical and cognitive) and extrinsic (socio-economic and environmental) factors [30]. The identification of risk factors for falls among older adults living at home has been the subject of several studies [31–34]. They have shown that the risk of falling increases as the number of risk factors increases.

3. Material and Methods

3.1. Approach for Analyzing and Improving Innovative Solutions

This article's assumption is that the notion of risk can contribute to the understanding of a problem situation and impact upon the design of innovative solutions. It may enrich the process of needs analysis as well as the evaluation and improvement of proposed solutions. The approach is depicted in Figure 2.

In the proposed approach, three stages of the design process are considered. The first step is the exploration of the problem. At this stage, different stakeholders (funders, project managers, future users, regulatory institutes) are involved in the needs analysis. Using different engineering and design tools, risks, lacks, dissatisfactions and desires are explored to analyze needs. Thus, the design objective and specifications can be defined to guide the design team in proposing solutions.

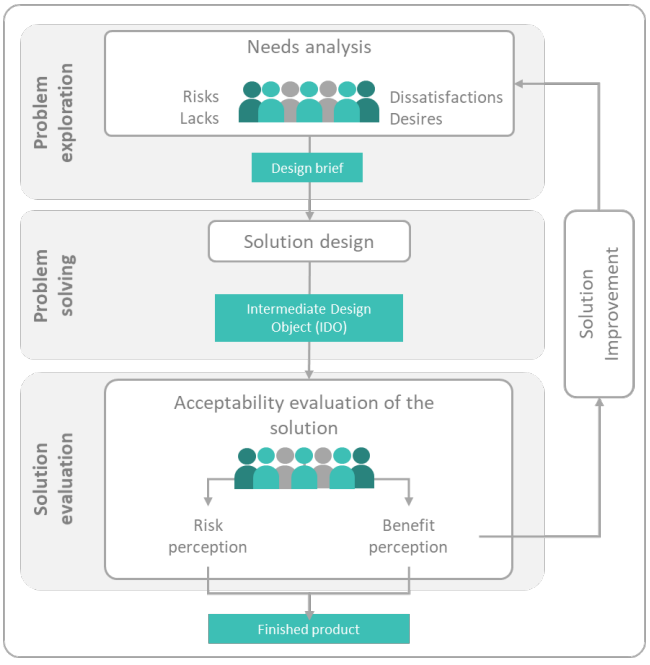


Figure 2: Approach proposed to analyse and improve an innovative solution.

The second step is the resolution of the identified problem. Here, different methods are used for the ideation and prototyping of solutions. In the first step, multiple creativity sessions are carried out in order to outline the design possibilities. Subsequently, an alternative solution is chosen and is then further designed. Different intermediate design objects (ICOs) are developed to communicate and validate the characteristics assigned to the solution. ICO is a partial outcome of the project, such as a schematic, a scale model, or a prototype [35]. The objective is to represent a part or the whole of the proposal in order to evaluate it with users and further develop the product.

The third step, solution evaluation, is an important stage in the design process. It allows the solution concept to be tested with future users to identify the benefits and risks they perceive relative to using this concept. In this way, the acceptability of a solution can be estimated. The result of this evaluation brings elements of improvement (to reduce the perceived risks and/or increase the benefits), which can be considered in the problem understanding and in the specification redefinition. This advanced understanding of the problem will allow designers to refine their proposals and better respond to it. We believe that successive evaluations of the solution and integration of gradual improvements would lead to an increase in acceptability. Thus, at the end of the design phase, the final product would have a higher probability of market acceptance.

The next section will apply this approach to the context of home care for the elderly. When we explored the problem of loss of autonomy among the elderly, the choice was made to apply the approach to the risk of falling. This risk has a high probability of occurrence and its consequences are severe.

3.2. Model for Risk Assessment Using Bayesian Networks – Risk of Falling

In order to estimate the probability of elderly people falling at home, we looked for a useful tool to assess the influence between factors and to determine the probability of occurrence. In the scientific literature, statistical methods are often used; however, they are insufficient for the evaluation of imprecise data from various sources. Recently, [36] used the Bayesian network modelling technique to estimate the probability of innovation acceptance based on an analysis of the influence of various factors. Among existing data modelling and analysis techniques, Bayesian networks are the most suitable for knowledge acquisition, representation and manipulation [37]. This technique makes it possible to reason about a problem by evaluating the causality between factors and sub-factors. In addition, it can be used to process uncertain information from different sources: expert knowledge, questionnaire data or the literature.

Bayesian networks have already been used for fall risk analysis [38, 39]. However, the work carried out to date has not sought to be part of an approach to design and improve innovative solutions.

This article proposes to use this technique to assess the probability of elderly people falling at home. It is also assumed that Bayesian networks can be used to make a wider and overall risk assessment of this population.

The construction of the Bayesian network for estimating the risk of falling was carried out following the steps proposed by

[37]. The first step consists in identifying the risk factors and the statuses that each of them may have. The second is the definition of the Bayesian network structure and the third is the construction of conditional probability tables.

3.3. Identification of Risk Factors

In the first stage, factors were identified mainly through the literature on the risk of falling [30-32,34,40-42]. A list was prepared and presented to an expert gerontologist and two researchers who analyzed and classified these factors (Table 1). Next, two possible statuses were defined for each factor: "yes" and "no". These statuses represent the variables of interest, in other words, the variations that each of the factors may present.

Table 1: Fall Risk Factor Codes

Code	Factor description
AG	Age
GD	Gender
HF	Fall historic
PP	Physiological problems
PP1	Balance disorders
PP2	Visual impairment
PP3	Orthostatic hypotension
CI	Cognitive impairment
CI1	Syncope
CI2	Dementia
CI3	Depression
SI	Social isolation
SI1	Loss of social and family ties
UL	Unhealthy lifestyle
UL1	Alcohol consumption
UL2	Malnutrition
UL3	Lack of physical activity
IE	Inadequate environment
IE1	Inadequate objects
IE2	Inadequate configuration of the living environment
MP	Medication problem
MP1	Poly medication
MP2	Taking psychotropic drugs

3.3.1. Definition of the Bayesian Network Structure

To build the structure of the network, a first proposal of links between factors was made based on the literature consulted. This structure was evaluated and improved with the experts. The final structure is shown in Figure 3. The network is made up of nodes, which correspond to the different factors and sub-factors. These nodes are linked by arcs that represent the causal and conditional dependence between them and the risk to be assessed. The factors have been classified into six groups: physiological problems, cognitive impairment, social isolation, unhealthy lifestyles, inadequate environment and medication problems. Each group is linked to between two and six factors. Four of these factors are related to more than one group: age, gender, history of falls and social isolation.

3.3.2. Construction of Conditional Probability Tables

Conditional probability tables are used to estimate the occurrence probability of one of the node statuses as a function of the parent nodes statuses [37]. A first version of the tables has been proposed from the literature consulted. The information obtained was supplemented by the experts, who used their experience and knowledge to define probability values between factors.

The final version of the network structure and the conditional probability tables by risk factor were modeled using the GeNIe Academic software.

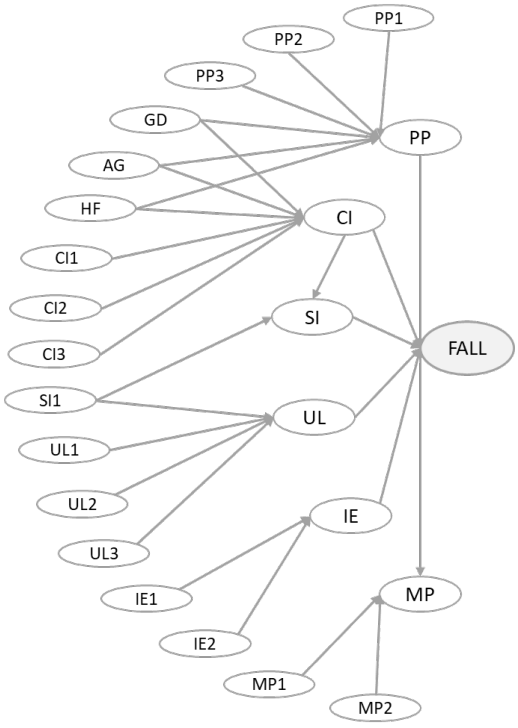


Figure 3: Bayesian network model of home falls risk

3.4. Data Collection Method

In order to make a first assessment of the probability of falling among older people, the proposed model was used. The data collection method chosen was the semi-directive interview. Interviews are one of the most commonly used methods in field studies to collect information about people's perceptions and preferences in light of their experiences [43]. The semi-directive interview serves to guide the discussion with the interviewee based on a list of themes rather than on direct questions [44].

A semi-directive interview guide was constructed. Following the recommendations of [44], the guide was composed of a list of themes related to the perception of exposure to the different risk factors of the model.

An older couple living independently in their home was selected for interview. Ms. M, 91 years old, and her husband Mr. A, 95 years old, have been married for 70 years. They have been retired for more than 30 years and wish to continue to live in their own home. These people were chosen because of their advanced age and their interest in continuing to live autonomously and safely in their home.

With the consent of these two persons and their children, the interviews were conducted at home. Two members of the research group participated in the development of the interview. One of the researchers led the discussion with each person, while the second researcher wrote down his observations on the development of the interview and on the characteristics of the household. The interviews were recorded and subsequently transcribed.

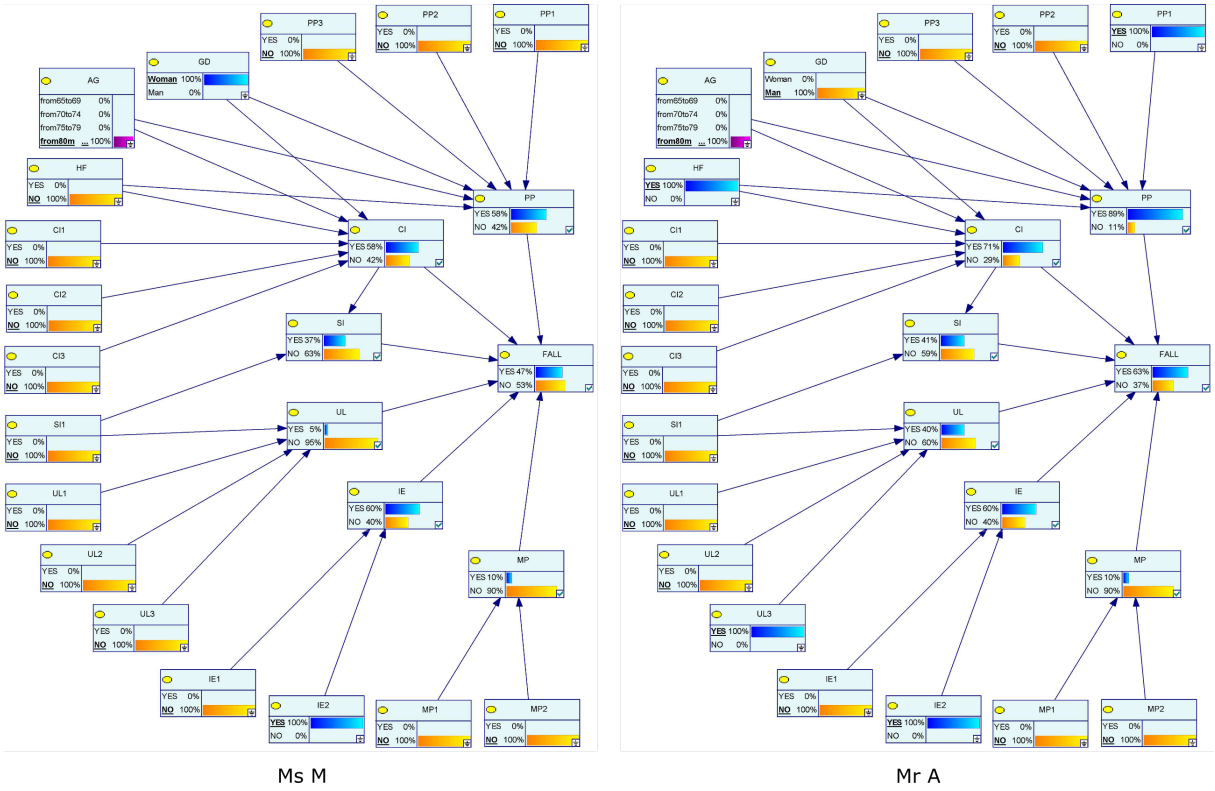


Figure 4: Fall risk estimation for Ms. M and Mr. A.

These elderly person's perception of the absence or presence of each of the risk factors was interpreted and coded according to the two conditions designated for each factor, namely, yes or no. The results obtained were introduced into the proposed Bayesian network.

The results obtained and the proposed Bayesian network were subsequently used to explore their use in the design and improvement of an innovative solution.

4. Results

4.1. Fall Risk Assessment in the Problem Exploration

In the interview, Ms. M declared that she has no physiological or cognitive problems. She does not consider herself to be in a situation of isolation since she has a large family and maintains strong relationships with them. She says that she does not have any health problems because she eats well, is very active and does not drink alcoholic beverages. In addition, she has home help to prepare meals, clean and care for her. She does not use psychotropic drugs and she does not take multiple medication.

Mr. A, in contrast to Ms. M, was to the victim of a serious fall in the house's bathroom a few months before the study. After this fall, several modifications were made to the bathroom by the family. However, Mr. A says he has difficulty moving around the house due to the "fear of falling again". This situation has increased his balance problems and decreased his physical activity.

In the observation we remarked that in fact several devices have been installed in the bathroom and in other parts of the house, such as in the bedroom and on the stairs. However, there are areas that may represent a risk: the floors are not non-slip in the kitchen, corridors and bedroom. Bathroom doors open inwards and locks cannot be opened from the outside. There are no brackets or grab bars in the corridors. They have a remote assistance system, which is installed in the master bedroom, but it cannot be used in the rest of the house.

The information obtained for Ms. M and Mr. A was synthesized in Table 2. Subsequently, this was entered into the Bayesian network and the probability of each person falling was estimated.

Figure 4 shows the networks per person. The existence of a risk factor is represented by the color blue and its absence by the color yellow. For the "age" factor, four bands were used, each represented by a different color.

Table 2. Response by factor

	Ms. M AG: 91 years old GD: woman	Mr. A AG: 95 years old GD: man
HF	No	Yes
PP1	No	Yes
PP2	No	No
PP3	No	No
CI1	No	No
CI2	No	No
CI3	No	No
SI1	No	No
UL1	No	No
UL2	No	No
UL3	No	Yes
IE1	No	No
IE2	Yes	Yes
MP1	No	No
MP2	No	No

In Ms. M's network, we can identify that the biggest risk factors are inadequate environment (60%), physiological disorders (58%) and cognitive impairment (58%), followed by social isolation (37%). The two other factors, lack of hygiene (5%) and medical problems (10%) are not as representative. Analysis of these factors yields a 47% probability of falling.

In Mr. A's network, the results are different. The risk factors for physiological (89%) and cognitive (71%) problems are very significant. This significance is probably a consequence of the fall he had previously experienced, which caused him balance problems. It may also have been due to the likelihood of isolation risk (41%) and poor hygiene (40%). The occurrence probability of the risk factors "unsuitable environment" (60%) and "medication problems" (10%) were no different from Ms. M's results since they live in the same house. Analysis of these factors results in a greater likelihood of falling, corresponding to 63%.

4.2. Design and Improvement of Innovative Solutions

Using Bayesian networks, the proposed model estimated the probability of two elderly people falling in their homes. An extended model could be useful to detect different risks. On the one hand, the identification of the risk of falling may help project managers to better understand the needs of this population and to define specifications for proposing solutions. On the other hand, based on the risk factors identified, a design team can define areas for improvement of an innovative design solution. These two perspectives for using the results will be explored below.

4.2.1. Designing Innovative Solutions

A variety of solutions can be proposed to reduce the probability of the risk of falling. Some proposals may focus on increasing the person's mobility capabilities according to the characteristics of their home. Other solutions could consist in adapting the living space to the person's abilities.

In the case of Mr. A, who has a high risk of falling, the biggest risk factors are: walking and/or balance disorders, poor configuration of the living space and lack of activity. We imagined a new concept of an exoskeleton that adapts to the person's morphology to help him or her move around at home. In order to find out whether this technology could be accepted by its future users, the perceived risks could be evaluated.

4.2.2. Improvement of an Innovative Design Solution

If an innovative solution is not acceptable, the risks identified could provide insights for improvement. For example, if we take the exoskeleton and assume that future users do not accept this technology, then based on the analysis of the risk of falling, some changes could be made. The impact of these improvements on the reduction of the fall probability can be estimated with the Bayesian network.

Table 3 summarizes the different fall probabilities for Ms. M and Mr. A when introducing the exoskeleton and incorporating modifications to it.

For Mr A, the introduction of the exoskeleton could reduce the probability of falling by 4%. However, if improvements are made to the solution to adapt it to the environmental conditions, such as a contact surface adaptable to floor characteristics, then the risk of falling could decrease to 54%. In addition, if the exoskeleton incorporates a robotic system to help maintain balance, change posture and move around, then Mr. A's level of physical activity will improve and, as a result, the probability of falling can be reduced to 50%. These improvements could reduce not only the risk of falling but also general risk perception. As a result, the acceptability of exoskeletons may increase.

Table 3. The probability of falling after improvements

	Fall probabilities for Mr. A	Fall probabilities for Ms. M
Initial situation	63%	47%
Exoskeleton	59%	43%
Exoskeleton with adaptable contact surface	56%	
Exoskeleton with robotic system	54%	
Exoskeleton with adaptable contact surface and robotic system	50%	

In the case of Ms. M, the exoskeleton could reduce the risk of falling by approximately 4%. The various changes made to this solution would not contribute significantly to a further reduction of this risk.

5. Discussion

A theoretical approach is proposed for the design, evaluation and improvement of innovative solutions. Unlike other existing methods [12], our proposal focuses on the notion of risk. In representing the approach, we have established three steps for the design process: problem exploration, problem solving, and solution evaluation. In the first phase, we propose that the needs analysis integrates the identification of risks in the initial situation. This would help the design team to identify needs that have not been expressed by stakeholders. In the second phase, these risks, translated into requirements, allow the designers to propose alternative solutions and develop different intermediate design objects. Finally, in the third stage, we propose to evaluate risk perception at the same time as benefit perception in order to estimate the acceptability of the solution. We believe that the results of this evaluation may be useful for a better understanding of the situation of the problem and for improving a solution being designed. Multiple iterations between problem understanding and solution assessment could lead to a product design that will better meet the needs of the target population and be more acceptable to the market.

As part of this approach, the work carried out focused on the risks for elderly people at home.

Risk identification for a person at home implies knowing the causal factors of each risk and relating them to each other. In order to propose a model for estimating the occurrence probability of a risk, we chose one of the most recurrent and significant risks, that of falling. This risk has already been studied [31-34], but no assessment model involving various types of causal factors has been proposed.

We proposed a model for estimating the risk of falling using the Bayesian network technique. Arbelaez already used this method to evaluate the acceptability of innovation [36]. Other

studies have used the same technique to assess the risk of falling [38, 39]. We use this technique to estimate the risk of falling by evaluating different types of causal factors: physiological problems, cognitive impairments, social isolation, environmental maladjustment, poor living conditions and medical problems. The results are intended to contribute to the analysis of needs and the evaluation and improvement of new solutions.

We used this model with two elderly people living at home. Interviews were conducted to obtain information. Using the proposed model, we identified a probability of falling for these two individuals. The results are different for each of them. One of the differentiating factors is the fact that one of these two people has already had a fall, which has affected their mobility and lifestyle.

The risk of falling is only one of the twenty-two identified hazards that can affect the ability to keep elderly people at home. The reasoning used to assess the risk of falling can be replicated to estimate other hazards. To help older adults live well at home, health innovations aimed at reducing risk are needed.

In view of the results obtained, we have imagined the type of solutions that could be proposed to reduce this risk. In addition, in designing an innovative solution, we hypothesized the improvements that could be made to a proposal, based on the risk of falling assessment.

6. Conclusion and Perspectives

This study was conducted with the aim to proposing a new design approach to innovative solutions based on a risk analysis of the real situation of elderly people at home.

In addition to this approach, a risk identification model is proposed, with the purpose of estimating the probability of a person falling. This work is evaluated with a case study concerning the design of an innovative solution for mobility assistance.

The Bayesian network technique was used to build the model. The causal factors of this risk were identified and correlated to define the structure of the model. We believe that this technique could be used for the evaluation of overall risks. This should be validated in future research.

The model developed was applied with two elderly people who were interviewed and observations were made in their homes. The responses obtained were used to feed the network and to determine the probability of each person falling. Based on the high risk of falling presented by one of the people studied, an exploration of the use of these results for the design and improvement of the acceptability of the solution was conducted.

The results presented here are a first exploration of the use of the proposed model. It is important for future studies to validate this approach with a larger number of elderly people.

Further work should consist in assessing risks and benefits globally and with other stakeholders (medical staff, family members, organizations). This evaluation should focus on the development of a fall prevention solution.

When assessing the perception of benefits and risks by multiple actors, various conflicts may arise. It is important to identify and address them. Managing these conflicts could increase the acceptability of an innovative solution.

The method used to build the Bayesian network can be improved. For the construction of the probability tables, several status spaces by factor could be considered in order to obtain more precise results, rather than only yes and no. In addition, field data could be used to complement expert opinions for the definition of the probability tables in a more detailed way.

In the context of a collaborative design approach for innovations in home support for the elderly, this model can be used as a tool for understanding the problem and assisting in decision-making. It will enable project leaders to evaluate and improve the acceptability of a solution, right from the early stages of design.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] N. Blanpain and G. Buisson, "Projections de population à l'horizon 2070, Deux fois plus de personne de 75 ans ou plus qu'en 2013," Paris, France, 2016.

[2] INSEE, "L'évolution de la dépendance des personnes âgées : un défi en termes de prise en charge et d'emploi", Orléans, 2012.

[3] S. Peek, "Understanding technology acceptance by older adults who are aging in place: a dynamic perspective," Ph.D Thesis, Universiteit op woensdag, 2017.

[4] F. Prate, "Evaluation des technologies pour la santé et l'autonomie : Vers un modèle transversal et modulaire. Illustration par la technologie robotique," Ph.D Thesis, Université de Nice Sophia Antipolis, 2015.

[5] A. M. Sponselee, B. Schouten, D. Bouwhuis, and C. Willems, "Smart home technology for the elderly: perceptions of multidisciplinary stakeholders," in Constructing Ambient Intelligence: AMI 2007 Workshops, Darmstadt, Germany, 2007. https://doi.org/10.1007/978-3-540-85379-4_37

[6] G. Arbelaez Garces, E. Bonjour, and A. Rakotondranaivo, "Towards a method for acceptability analysis: Application to healthcare innovation," in International Conference on Engineering, Technology and Innovation: Engineering Responsible Innovation in Products and Services (ICE 2014), Bergamo, Italy, 2014. https://doi.org/10.1109/ICE.2014.6871583

[7] L. Acosta-Salgado, A. Rakotondranaivo, and E. Bonjour, "Innovation acceptability in elderly care: a risk evaluation approach," in IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Sophia-Antipolis, France, 2019. https://doi.org/10.1109/ice.2019.8792564

[8] S. Cleary and T. Malleret, Risques. perception, évaluation, gestion : Une approche positive des risques globaux auxquels sont confrontés les décideurs. Maxima, 2006.

[9] L. Acosta-Salgado, L. Morel, and I. Verilhac, "Towards a better understanding of the concept of design," Proj. / Proyéctica / Proj., 2(20), 91–114, 2018. https://doi.org/10.3917/proj.020.0091

[10] N. Skiba, "Processus d'innovation centré sur l'utilisateur: identification des besoins et interprétation de données issues de l'intégration de l'utilisateur dans le processus de co-conception," Ph.D Thesis, Université de Lorraine, 2014.

[11] Carine Lallemand and G. Gronier, Méthodes de design UX, EYROLLES, 2016.

[12] H. BEN Rejeb, V. Boly, and L. Morel, "A new methodology based on kano model for the evaluation of a new product acceptability during the front-end phases," in IEEE International Computer Software and Applications Conference, Turku, Finland, 2008. https://doi.org/10.1109/COMPSAC.2008.94

[13] M. Fishbein and I. Ajzen, Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research, Addison-We, 1975.

[14] I. Ajzen, "The Theory of Planned Behavior," Organ. Behav. Hum. Decis. Process., 50(2), 179–211, 1991. https://doi.org/10.1016/0749-5978(91)90020-T

[15] F. D. Davis, "A technology acceptance model for empirically testing new end-user information systems: Theory and results," Ph.D Thesis, Massachusetts Institute of Technology, 1986.

[16] H. Taherdoost, "A review of technology acceptance and adoption models and theories," in 11th International Conference Interdisciplinarity in Engineering, INTER-ENG 2017, Tirgu Mures, Romania, 2017.

https://doi.org/10.1016/j.promfg.2018.03.137
[17] Y. Tanaka, M. Kitayama, S. Arai, and Y. Matsushima, "Major psychological factors affecting consumer's acceptance of food additives: Validity of a new psychological model," Br. Food J., 117(11), 2788–2800, 2015. https://doi.org/10.1108/BFJ-02-2015-0062
[18] A. Bearth, M. E. Cousin, and M. Siegrist, "The consumer's perception of artificial food additives: Influences on acceptance, risk and benefit perceptions," Food Qual. Prefer., 38, 14–23, 2014. https://doi.org/10.1016/j.foodqual.2014.05.008
[19] J. Choi, A. Lee, and C. Ok, "The effects of consumers' perceived risk and benefit on attitude and behavioral intention: a study of street food," J. Travel Tour. Mark., 30(3), 222–237, 2013. https://doi.org/10.1080/10548408.2013.774916
[20] M. S. Featherman and P. A. Pavlou, "Predicting e-services adoption: A perceived risk facets perspective," Int. J. Hum. Comput. Stud., 59(4), 451–474, 2003. https://doi.org/10.1016/S1071-5819(03)00111-3
[21] D. Delignières, "Risque préférentiel, risque perçu et prise de risque," In : Cognition et performance [online], Paris, France, 1993. https://doi.org/10.4000/books.insep.1419.
[22] S. Vandentorren et al., "August 2003 heat wave in France: Risk factors for death of elderly people living at home," Eur. J. Public Health, 16(6), 583–591, 2006. https://doi.org/10.1093/eurpub/ckl063
[23] M. D. L. Á. Vázquez-Sánchez, M. D. C. Gastelu-Cantero, and J. L. Casals-Sánchez, "Valoración de las necesidades de los ancianos que viven solos en una zona básica de salud," Enferm. Clin., 18(2), 59–63, 2008. https://doi.org/10.1016/S1130-8621(08)70699-5
[24] S. e. Carter, E. m. Campbell, R. w. Sanson-Fisher, S. Redman, and W. j. Gillespie, "Environmental hazards in the homes of older people," Age Ageing, 26(3), 195–202, 1997. https://doi.org/ 10.1093/ageing/26.3.195
[25] A. E. Stuck, J. M. Walthert, T. Nikolaus, C. J. Büla, C. Hohmann, and J. C. Beck, "Risk factors for functional status decline in community- living elderly people: a systematic literature review," Soc. Sci. Med., 48(4), 445–469, 1999. https://doi.org/ 10.1016/s0277-9536(98)00370-0
[26] C. Berr, F. Balard, H. Blain, and J. Robine, "Vieillissement l'émergence d'une nouvelle population," Med. Sci., 28(3), 281–287, 2012. https://doi.org/10.1051/medsci/2012283016
[27] H. J. Swift, D. Abrams, R. A. Lamont, and L. Drury, "The risks of ageism model: how ageism and negative attitudes toward age can be a barrier to active aging," Soc. Issues Policy Rev., 11(1), 195–231, 2017. https://doi.org/10.1111/sipr.12031
[28] S. Clement, C. Rolland, and C. Thoer-Fabre, Usage, normes, autonomie : Analyse critique de la bibliographie concernant le vieillissement de la population, Puca, 2007.
[29] M. E. Tinetti, C. Gordon, E. Sogolow, P. Lapin, and E. H. Bradley, "Fall-risk evaluation and management: challenges in adopting geriatric care practices," Gerontologist, 46(6), 717–725, 2006. https://doi.org/10.1093/geront/46.6.717
[30] M.-L. Gaubert-Dahan, A. Cougnaud-petit, L. De Decker, C. Annweiler, O. Beauchet, and G. Berrut, "Essai de modélisation des facteurs de risque de chute chez les sujets âgés," Ger Psychol Neuropsychiatr Vieil, 9(3), 277–285, 2011. https://doi.org/10.1684/pnv.2011.0289
[31] R. Rajagopalan, I. Litvan, and T. P. Jung, "Fall prediction and prevention systems: Recent trends, challenges, and future research directions," Sensors, 17(11), 1–17, 2017. https://doi.org/10.3390/s17112509
[32] C. A. Pfortmueller, G. Lindner, and A. K. Exadaktylos, "Reducing fall risk in the elderly: risk factors and fall prevention, a systematic review," Minerva Med., 105(4), 275–281, 2014.
[33] B. Leclerc et al., "Risk factors for falling among community-dwelling seniors using home-care services : An extended hazards model with time-dependent covariates and multiple events," Chronic Dis. Can., 28(4), 111–120, 2008.
[34] L. Z. Rubenstein, "Falls in older people: epidemiology, risk factors and strategies for prevention," Age Ageing, 35(2), 37–41, 2006. https://doi.org/10.1093/ageing/afl084
[35] D. Galvez Manriquez, "Evaluation de la capacité à innover : Une approche par auto évaluation et suivi supporté par des analyses multicritères dynamiques," Ph.D Thesis, Université de Lorraine, 2015.
[36] G. Arbelaez Garces, A. Rakotondranaivo, and E. Bonjour, "An acceptability estimation and analysis methodology based on Bayesian networks," Int. J. Ind. Ergon., 53, 245–256, 2016. https://doi.org/10.1016/j.ergon.2016.02.005
[37] P. Naïm, P.-H. Wuillemin, P. Leray, O. Pourret, and A. Becker, Réseaux Bayésiens, 3e edition, Eyrolles, 2007.
[38] G. Koshmak, M. Linden, and A. Loutfi, "Dynamic Bayesian networks for context-aware fall risk assessment," Sensors, 14(5), 9330–9348, 2014. https://doi.org/ 10.3390/s140509330
[39] H. Wang, M. Li, J. Li, J. Cao, and Z. Wang, "An improved fall detection approach for elderly people based on feature weight and Bayesian classification," in 2016 IEEE/ICMA International Conference on Mechatronics and Automation, 2016. https://doi.org/10.1109/ICMA.2016.7558609
[40] American Geriatrics Society, "Guideline for the Prevention of Falls in Older Persons," *J. Am. Geriatr. Soc*., 49(5), 664–672, 2002. https://doi.org/10.1046/j.1532-5415.2001.49115.x
[41] P. J. Fos and C. L. McLin, "The risk of falling in the elderly: A subjective approach," Med. Decis. Mak., 10(3), 195–200, 1990. https://doi.org/10.1177/0272989X9001000306
[42] HAS (Haute Autorité de Santé), "Évaluation et prise en charge des personnes âgées faisant des chutes répétées," Paris, 2009.
[43] S. Q. Qu and J. Dumay, "The qualitative research interview", 8(3), 2011. https://doi.org/ 10.1108/11766091111162070
[44] L. Garreau and P. Romelaer, Méthodes de recherche qualitatives innovantes. Economica, 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 159-168 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Testing Web Service Compositions: Approaches, Methodology and Automation

Dessislava Petrova-Antonova[1,*], Denitsa Manova[2], Sylvia Ilieva[1]

[1]*Software engineering, Sofia University, 1113, Bulgaria*

[2]*Rila Solutions, 1113, Bulgaria*

ARTICLE INFO

Article history:
Received: 05 November, 2019
Accepted: 20 December, 2019
Online: 22 January, 2020

Keywords:
Testing approaches
Testing methodology
Testing web service compositions

ABSTRACT

Web services give a new view of the web as the biggest, widely accepted and the most straightforward distributed software platform. Their composition into applications and business processes is still a complex, non-trivial task, requiring highly rational efforts not only from the software developers, but from the quality assurance specialists. The provision of web service compositions' quality brings a lot of challenges due to variability of difficulties at infrastructure, service and choreography levels and the need of different types of testing in unknown context and environment. A consolidated quality assurance methodology that advances the fundamental understanding of testing in terms of concepts, models, techniques, standards and automation is required. This methodology needs to enable effective exploration, comparison, evaluation and selection of testing approaches, platforms and tools.

This article proposes such a methodology and reviews the current testing approaches for single and composite web services from an objective, holistic perspective. The methodology is presented as an end-to-end testing procedure, in which activities are facilitated by a set of testing approaches, techniques and best practices. A concrete solution is recommended for actual implementation of each activity either through selection among the most appropriate and effective existing approaches or development of new approaches, mainly in case of critical issues such as dependencies analysis, partner web services' isolation and injection of faults. A common framework that integrates different testing tools automates the methodology. Its applicability, completeness, level of automation, and level of novelty is evaluated through testing of real testing scenarios.

1. Introduction

Web services provide a novel paradigm for interoperability of distributed applications. They act as collaborative agents to deliver advanced and high-quality functionality to the end users. That is why the web services are used by the successful enterprises to build flexible and fast connections with their partners, and thus reducing the cost and increasing the revenue. Such business interactions are possible through implementation of Web Service Compositions (WSCs) in a standardized, secure and interoperable manner. Although the WSCs provide a lot of benefits for the developers, they bring challenges to testing as well as to overall quality assurance process. First, its implementation is often based on web services, which are developed and deployed on diverse environments supported by different vendors. In case of legacy systems and components, the web services act as wrappers of functionality that is hard to be controlled and simulated in a testing environment. The orchestrated web services are not always available or could be undeployed by the provider, which leads to additional efforts for emulation of their behavior during testing. In addition, the emulation is required in case of payed web services to reduce the testing process's cost. Appropriate message data needs to be generated not only for the emulated web services to mimic their behavior, but for the composition as whole in order to be invoked in the testing environment. Achieving a high level of test coverage and determination of failure causes is difficult due to the missing programming code of the orchestrated services as well as the requirement for creating test cases based only on their public interfaces. Such interfaces are usually defined using the Web Service Description Language (WSDL). When a dynamic binding

[*]Dessislava Petrova-Antonova, 1113 Sofia, Bulgaria, 125 Tsarigradsko shose Blvd. Blok 2, +359887572094 & d.petrova@fmi.uni-sofia.bg

of web services is implemented, the emulation process becomes more complex, since advanced stubs (mockups) should be developed.

A lot of testing approaches and automated tools have been proposed to meet the above challenges. Unfortunately, the most of them cover only single testing activities such as analysis of testing paths, generation of test cases, web service mocking, injection of faults, etc. However, it is important to merge all testing activities in a general testing procedure supported by end-to-end methodology and thus to facilitate the whole testing process. This article is an extension of work originally presented in International Conference on the Quality of Information and Communications Technology (QUATIC'18) [1], which proposes such a methodology, called Testing As Service Software Architecture (TASSA). The methodology addresses the following major problems of WSC testing:

- Difficulties in emulation of web services deployed on heterogeneous platform and controlled by external providers;
- Missing or temporally unavailable web services and inability to detect the causes of failures;
- Lack of automation to test unanticipated behavior of web services and test case generation to reach a high level of test coverage.

Therefore, the benefit of the proposed methodology are as follows:

- Provision of end-to-end testing methodology for WSC;
- Recommendation of concrete approaches for actual implementation of each activity of the methodology;
- Development of completely new approaches to cover critical issues such as analysis of data dependencies, isolation of partner web services and injection of faults;
- Full automation in a single framework for testing WSCs, described with Business process execution language (WS-BPEL) [2].

The next two sections present the TASSA methodology and the testing approaches appropriate for its implementation. Section 4 provides a view of TASSA methodology's validation and Section 5 summarizes the validation results. Section 6 concludes the paper.

2. TASSA methodology

TASSA methodology supports functional, performance and security testing of WSCs and provide means for validation of their behavior if implementation changes are in place. It relies on a small number of artefacts, such as a BPEL file, input data, expected output data and test assertions. It consists of seven main activities, presented with a workflow diagram in Figure 1.

2.1. Prerequisites

W3C defines a set of standards for WSCs that are used for standard compliance validation of the WSCs [3]. If some inconsistencies are detected, a notification to perform the approapriate coreections are sent to the tester. The isolation of a partner web service can can follow two approaches: (1) invocation of mockup service that mimic the behavior of the partner web service or (2) generation of expected response that is expected from the partner web service. TASSA methodology adopts the second approach based on change of the communication channel between the WSC and its partner web service. This is realized using a so called "mediator service" allows fo injection of different types of faults in the communication channel.

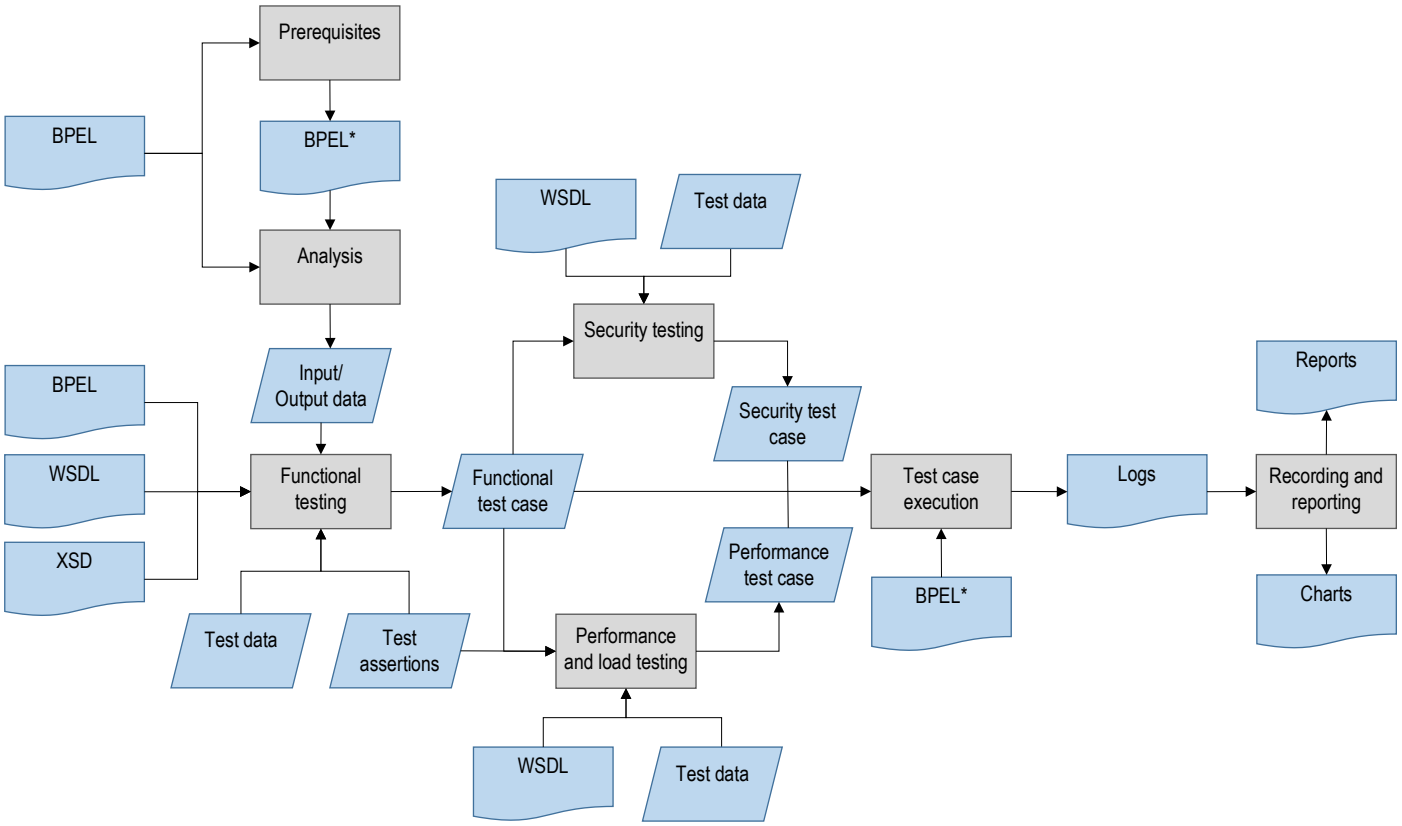


Figure 1: Methodology workflow

2.2. Analysis

In order to obtain all test paths, the WSC is usually represented using a specific formal model. TASSA methodology transforms the WSC in a BPEL Flow Graph (BFG) similarly to the approaches in [4] and [5]. First, all paths in the BFG are identified, next the unfeasible ones are filtered and the finally the feasible paths are used for generation of test cases.

2.3. Functional testing

The WSDL description of the WSC is the primary source for functional test case generation. The main components of a given test case are the request with input data and the expected response with output data. Similar test cases are a common practice for testing single web services. A modified approach that support tracing of the executed activities in the WSC is applied by the TASSA methodology.

The generation of test data is based on a widely adopted approach based on XML Schema Definition (XSD) of the WSC. It is further developed to support parametrization of test cases using variables, functions as well as connectivity with external data pools allowing for data-driven testing. The expected results from the execution of the WSC are a base for definition of test assertions. The execution order of the activities in the WSC is an important assertion, when the behavior of multiple web services is orchestrated. Additional assertions relevant to both single and composite web services are defined, based on a single value, a value type, a regular expression, an HTTP status code, a file type and size, etc.

2.4. Security testing

The request to the WSC is enriched with appropriate test artefacts for the purposes of security testing. The response from the WSC is checked whether it matches the security requirements, such as delivery of partial message, unauthorized access, unavailable partner web service, etc. At the message level, the security testing includes sending of "secure" request and checking the response sent by the web service. There are a lot of security mechanisms applicable to WSCs. The TASSA methodology proposes usage of WS-Security standard and QoS policies, which are based on enrichment of exchanged messages with a specific information based on WS-RM, WS-Addressing or MOTM. At transport level, the testing for security is related to testing of the authentication, including sending certificates like Kerberos, SSL, etc.

The TASSA methodology follows a new approach for the robustness testing. It transforms the WSC under test allowing for simulation of faults during its execution. The supported faults are an unavailable partner web service, a delay of the response from a partner web service, wrong structure or semantic of the response from a partner web service. The test assertions for security testing are defined in a similar way of functional testing.

2.5. Performance and load testing

The performance and load testing validate the non-functional characteristics of the WSC. Since, it requires a set of requests to be send in order to simulate parallel executions, a definition of virtual users is performed for single functional test case or a group of them for a given time interval. The threads are commonly used in the most cases. They are scheduled to define the way, on which the virtual users are managed. The schedule contains information about the time for starting and stopping of requests, execution order of requests and time interval between them, number of executions, etc.

The test assertions for performance testing are based on those for functional testing. Test assertions for time intervals, resource usage, file sizes and others characteristics are also defined. They provide insight for the system and network load, the hardware and software resources, etc. Additionally, a special assertion for checking the compliance with the Service Level Agreements (SLAs) is created.

2.6. Test case execution

Prior execution of functional test cases, the BPEL file of the WSC should be deployed on the application server and an instance of the WSC should be created. The request defined in the functional test case is sent to the WSC and its behavior is tested in a similar way to a single web service. The security testing requires to configure the security settings of the application server, where the BPEL file is deployed. The performance or load testing needs definition and management of threads according to the parameters of in the test cases. When a huge load needs to be simulated or hardware resources of the application server are limited, the execution of test cases is distributed among different machines.

2.7. Recording and reporting

The information obtained during the execution of the test cases is recorded in a log file. In order to allow for easily processing, the valuable outputs are collected in a test data storage. Apart from the directly obtained test results, additional metrics are recorded, which are related to the WSC itself, the application server, communication channels, the network and so on. These metrics are simple, but can be further aggregated for computing of more complex ones. In order to check whether the test is failed or passed, the expected output and actual one after test case execution are compared based on the test assertions. When the test is failed, the reason for the failure is collected for defect allocation as well as for a subsequent regression testing. The test reporting uses the data from the log files and the corresponding test data storage. The XML is the most commonly used data format for the test reports.

3. Testing Approaches

This section presents testing approaches that are appropriate for implementation of TASSA methodology. For each methodology's task several approaches are identified and the most appropriate ones are recommended. For particular tasks new approaches are developed and reasoning about their applicability is given.

3.1. Standard compliance checking

The preparation of WSC for testing includes three main activities, described in previews section, namely checking for compliance with standards, isolation of partner web services and injection of faults.

The compliance with the standards can be validated by the most of Integrated Development Environments (IDEs) such as

Eclipse BPEL Validator, Oracle ILINT and NetBeans XML Validate. It can be performed following three different techniques depending on the validation rules:

- Rules described with a programming code (e.g. code written with Java);
- Rules based on XML or OCL constraints;
- Rules defined with a model of the WSC description (e.g. UML model).

The proposed methodology relies on the Eclipse BPEL Validator. It processes the BPEL file as a DOM and for each node (assign, copy, from, to, etc.) creates rules for checking the constraints prescribed by the corresponding standard. The rules are written in Java and can be changed by the tester.

3.2. Isolation of partner web services

The replacement of partner web service with a stub (mockup) service is a technique widely used to remove external dependencies from the WSC [6,7]. An alternative approach is to generate a message that substitutes the response expected by the WSC from a partner web service [8,9,10]. The second approach requires the real communication with the partner web services to be recorded and appropriate interfaces and data to be manually defined.

For the purpose of TASSA methodology, a new approach is elaborated. It replaces the BPEL activities related to external dependencies (e.g. Invoke activities) with ones that are internal to the WSC and provide similar output to the original activities (e.g. Assign activities). Since the output of the simulated activities are known, the execution of the WSC can be controlled to follow a particular path. For this purpose, proper values for the variables in conditional branches should be defined. Identification of the sequence of branch conditions and the values for conditional variables is a heavily task but it is automated by the TASSA isolation tool.

3.3. Fault injection

The robustness of the WSC can be tested relying on methods for negative testing by causing faults in the WSC itself or in its communication with the partner web services. A classification of the faults that are typical for service-oriented systems can be found in [11], while the specific faults for WSC are described in [12]. In [13] authors propose an approach that injects faults in the partner web service and then perform coverage testing of the code responsible for recovery from failures (e.g. exception handlers). In [14] a fault injection technique is applied by exploring the behavior of the WSC regarding the injected faults and thus to assess the its fault tolerance capabilities.

The fault injection is the second new approach proposed by the TASSA methodology. It supports injection of delays, errors and other malfunctions in the message exchange between WSC and its partner web services. The BPEL file of the WSC is modified by replacing the call to a partner web service with a call to a proxy service. In case of faults, the behavior of the third-party services is out of the TASSA scope. Therefore, faults in the outgoing calls are not injected. The proposed approach requires the original partner web service to be called and then intervention in its response to be performed. For this reason, the proxy service calls the original partner web service and then using the information for the fault simulated performs the required action. When data errors need to be injected, the two options are available – to change the values of data fields, while keeping the structure of the message or to insert random errors in the message data, which could break the XML structure of the response. For the first case, a tool for random generation of data according to an XML schema is required.

3.4. Dependency analysis

The analysis of WSC is focused on states, transitions and related usage of BPEL activities that are part of complex interactions. The software systems support two kinds of interactions as follows:

- Data flow interactions, where the definition and values of the next data items depend on the definition and values of the previous ones.
- Control flow interactions along the execution path, in which the next executions depend on the previous ones;

The testing of the above interactions is covered by the data flow testing and control flow testing, respectively. Such techniques are applicable mainly to white-box testing. In the context of the service-oriented systems, there are additional interactions due to communication with external services, for which only the types of input and output data are known.

An overview of types of interactions, related dependencies and approaches for analysis is can be found in [15]. The current approaches for dependency analysis apply variety formal models. The most popular ones transform the WSC in a control flow graph [16,17], Petri net [18,19], a state graph [20] or an UML model [21,22].

For the purpose of TASSA methodology, a new approach for dependency analysis is developed. As was mentioned before, it transforms the WSC in a BFG in order to find all executable paths. The dependency analysis includes identification of the conditional statements together with their conditional variables and constants, which values guide the execution of the WSC on a certain path.

3.5. Test data generation

The test data generation is a complex task, since the programming code of the WSC itself and its partner web services is often unavailable. That is why the approaches for test data generation are “black box” based, meaning that they rely only on the scope and type of the input and output variables.

Since the communication in WSCs is performed through exchange of XML messages, a large group of test data generation techniques use WSDL descriptions of the web services in combination with their XSDs [23,24,25,26,27]. The XSDs determine the constraints over the data types that are useful for generation of both simple and complex test data.

The approach proposed for TASSA methodology is similar to those presented in [28,29,30,31], since it also processes XSDs. It produces XML instances from a given XML Schema available in the WSDL file or in the BPEL file. Thus, XML messages needed for communication of the WSC with its partner web services are generated. The proposed approach can be applied to functional and robustness testing of WSCs due to support of both correct and incorrect XML instances’ generation. The implemented algorithm supports generation of empty XML documents – XML instances, which structure follows a given XML Schema, but does not

contain concrete values for the XML elements. In addition, generation of XML instances populated with random valid and invalid data depending on the testing goal is provided. Finally, a manual specification of values for the XML elements or their loading from external file are supported.

3.6. Test assertions definition

The service-oriented systems require definition of test assertions for a large group of test artefacts such as file content, file properties, message content (both header and body), message properties, expressions (e.g. XPath), variable values, variable properties, time properties and sequences of activities' execution.

Durand and Green propose a test script model allowing for expression of predicates crossing over diverse inputs and to handle a richer spectrum of outcomes [32]. The test assertions are well-structured objects defined with identifier, source (normative conformance requirement), target (instance of a specific artifact type), pre-condition that should be fulfilled, predicate (logical expression over the target) and prescription level (level of imperativeness of the source requirement). The major benefit of such approach is that the target can be tested within its context. The test assertions are defined using XLST and XPath and can be applied to functional testing and non-functional testing.

The power of XPath functions and expressions is leverage by Schematron – a simple pattern and rule language suitable for document testing [33]. Since the handling of the namespace is difficult, when rules are written according to an XML schema, and when there is a namespace prefix in a value of an attribute or element, the language is primarily intended for XML instances. XMLUnit is a tool offering both Java and .NET interfaces, which also supports validation according to a given XML schema, assertion of the values of XPath queries or comparison of XML documents with the expected outcomes [34]. The approach behind the tool relies on JDOM to traverse the XML tree.

TASSA methodology adopts the approach proposed by Durand и Green. In addition to it, the XMLUnit is also utilized due to its high level of automation.

3.7. Functional test case generation

The approaches for generation of functional test cases can be divided in two groups depending on the type of testing technique – white box or black box. The black box techniques handle the WSC as a single service with known operations and message types defined in a WSDL document. On the other hand, the white box techniques explore the actual BPEL file to find information about activities, control flow and data flow. According to the formal model used for transformation of WSC to find executable paths and thus generate test cases, the approaches falls in three main groups: approaches based on CFG [16], approaches using Petri nets [18] and approaches adopting model checking technique such as Web Service Automata (WSA) [35], Stream X-machine (SXM) [36], UML 2.0 activity diagram [37], PROMELA [38], etc.

The solution for test case generation of TASSA methodology follows mixed approach – a combination of black box and with box techniques. It supports the following tasks: (1) determination of web service operations and BPEL variables; (2) generation of templates of SOAP request; (3) test assertions' definition at HTTP level, SOAP level and BPEL variable level; (4) execution of test cases by exchange of SOAP messages; (5) collection of test results for reporting and follow-up actions. The test cases are described and stored in an XML format, where the root element has two attributes – "name", corresponding to the test case name, and "template", indicating whether a data-driven testing can be performed. The other elements are a narrative description of the test case, a WSDL operation under test, an input data, a data source for the purpose of data-driven testing, and test assertions. The provided types of assertions verify the status code of the HTTP response and response time, validate the SOAP message in case of success or failure and can apply XPath expressions to a BPEL variable or to a SOAP body.

The test case provides information for the web service address, and the message parts like HTTP headers, authentication mechanisms if they are required, and data that is carried in the body of the SOAP request. The test result consists of HTTP headers, data that is carried in the body of the SOAP response, the BPEL variables' values and the execution time. In case of data-driven testing, parametrized test cases are generated, in which the actual data in the SOAP request body and the expected data in the SOAP response body are substituted with variables.

3.8. Performance testing

Since the web services are inherently concurrent, the concurrent issues continue to receive a huge interest. There are two primary solutions for implementation of concurrency models, namely event-based and thread-based. In thread-based applications each request is resolved in synchronous mode. Thus, each request is processed in a single thread. When the response is complete, the thread returns back to thread pool. If an external service is called to resolve the request, the thread has to wait for the response. In event-based applications each request is resolved in an asynchronous mode. The thread returns back to thread pool before the response is completed and it is ready to serve some other request. There is a single thread (or a small number of threads) that routes and manages all requests. If an external service is called to resolve the request, the process continue to execute without blocking and receives a response through a callback mechanism. The BPEL engines work following one of the two models. IBM WebSphere Process Server, ActiveBPEL Engine, Oracle BPEL Process Manager are thread-based, while BPWS4J and BPEL-Mora are event-based. A BPEL engine prototype with a high performance, proposed in [39], adopts and event-driven architecture and join patterns delivered by the Microsoft Concurrent Coordination Runtime (CCR). Apache ODE executes long-running business processes, described with BPEL. ODE's Java Concurrent Objects (JaCOb) framework ensures the runtime implementation of BPEL constructs at the instance level [40]. JaCOb supports a concurrency at application-level without relying on threads.

The type of concurrency model directly affects the performance of WSC. In addition to the Response time, other non-functional characteristics of web services related to their performance are Throughput, Availability, Accessibility and Successful rate. In the context of WSC, the performance depends on the orchestration of the partner web services, which can follow five patterns: sequence parallel, conditional, cyclic or discriminative execution. For each execution pattern, the quality factor of each partner service is calculated and then an aggregated factor for the whole WSC is delivered. The Response time, Availability, Reliability, Bandwidth and Cost are non-functional characteristics used for assessment of WSCs in [41]. A drawback of the proposed approach is that it considers only sequence

execution of partner web services. The approach presented in [42] covers all five execution patterns, but takes into account only three are non-functional characteristics (Response time, Reliability and Cost). In contrast to the approaches that apply aggregation method for quality assessment, there is a more realistic method, which accounts for the uncertainty of partner web services and estimate the quality of a WSC's probabilistically.

TASSA methodology proposes usage of combination of approaches. It is validated in the context of Apache ODE BPEL Engine. The generation of the virtual users us performed by external tool, which handles the WSC as a single service applying black box testing. The same number of requests following the same schedule are used for testing of WSC and its partner web services.

3.9. Security testing

The security testing of WSCs brings challenges due to their distributed nature and requires application of more than a single approach. Kabbani and Tilley consider the security of WSCs from three points of view: known threats and weaknesses, security mechanisms for web services and privacy, integrity and accessibility [43]. The Open Web Application Security Project (OWASP) works on improving the software security. It describes a set of security vulnerabilities, including those that affect the web services. It proposes tools and best practices for security assurance [44]. The OASIS consortium describes security quality factors such as Encryption, Authorization, Authentication, Non-Repudiation, Integrity, Availability, Privacy and Audit [45]. The WS-standards (WS- Security Policy, WS-Addressing, WS-SOAP Message Security, WS-Trust) prescribe mechanisms for secure communication between single web services, but they are applicable to WSCs too.

Several research works propose development of models for data access of single services [46,47,48]. The security assessment of WSCs is based on such models. An alternative approach is to use a web service mediator, called broker, which composes the web services against the specified security constraints. An optimal WSC is selected according the QoS properties [49]. The broker uses a repository to store security and quality properties of web services and updates these properties after each execution of WSC. The security certification of web services is a technique applicable to selection, discovery, and composition processes. It includes certificate issuing and management, certification-aware service discovery, certificate validation, and service consumption [50]. Biskup et. al propose a framework supporting a decentralized execution of WSCs, which guarantee the correctness and the security of the execution [51]. The approach is based on a container – an encrypted and authenticated data structure that is passed among the composed web services. An XML-based script language for definition of security policies of WSCs is presented in [52]. It is integrated to a platform for run-time monitoring and security analysis. Data sharing agreements are another solution for secure collaboration between parties [53]. They define the data sharing policies for authorizations, obligations and prohibitions. The policies indicate the authorized, obliged or denied actions together with their related data and subject.

TASSA methodology relies on validation according to the security standards, testing through fault injection, described in Section 3.3 and checking of the implemented security mechanisms. The quality factors of security in [45] are defined for single web services. The security tests check whether particular mechanisms and standards are applied or not. They can be performed on a transport level or on a message level. Since TASSA methodology are focused on WSCs, it proposes testing of the data exchange security, as it is prescribed in [53].

3.10. Recording and documentation of test results

The instrumentation is a popular method for monitoring and analysis of the execution of WSCs is [54,55,56]. It adds s specific code in the BPEL file, which collects execution information without changing the normal execution of the WSC. The collected information is sent to external auditing web service for processing. The approach is technology independent and is suitable for functional testing, since it allows monitoring of data flow and control flow. It can be applied to security testing, but it is not appropriate for performance testing due to invocation of external web service, which can slow down the performance.

The BPEL engine itself provides a detail information for the execution state of the WSC. For example, the ActiveBPEL Engine [57], Oracle BPEL Process Manager (Glassfish) [58], IBM WebSphere Process Server [59], Apache ODE [40] collect information about the Response time, the variables and activities as well as the invoked partner web services and the respective message exchange. Since the BPEL engines control the execution of WSC, the approaches and tools relying on their capabilities are technology dependent, but applicable to both functional and non-functional testing. SALMon is a service-oriented system for monitoring services in order to detect violations in service level agreements (SLAs) [60]. It is technology independent, including three types of components, namely Monitors, Analyzer and Decision maker. The Monitors consist of measuring instruments, the Analyzer checks the SLA rules, while the Decision maker performs corrections to satisfy SLA rules. An architecture for run-time monitoring of WSCs, described with BPEL, is proposed in [61]. The business logic of the composition and the monitoring functionality are clearly separated through implementation of two types of monitors – instance monitors and class monitors. The instance monitors are responsible for execution of a single instance of the composition, while the class monitors report aggregated information about all executed instances. Both monitors are specified through a specific language and are generated as Java programs, which can be deployed in the run-time environment of the monitor engine.

TASSA methodology adopts the approach relying on the BPEL engine for monitoring of WSC and subsequent reporting of the collected information. Such approach is suitable for all types of testing and provides information about assertions, variables, messages, control flow and communication with the partner web services. Its advantage is the lack of intervention in the WSC and early reporting of the test results.

4. Validation of TASSA methodology

The validation of the TASSA methodology is performed through its application for testing of a sample business process, called *Check Payment*. The testing is performed using NetBeans integrated development environment.

4.1. Functional testing

The definition of functional test cases includes the following steps:

1) Selection of BPEL file, containing description of the *Check Payment* business process;

2) Selection of test path from the data dependency tree, generated from the BPEL file;

3) Identification of variables and constants, from which the execution of the selected path depend on, using the data dependency analysis tool. The tool returns description of dependencies, shown in Listing 1;

Listing 1: Description of dependencies

```
This branch is ELSE
This branch is IF
This branch is IF Condition 0:
String representation:
(count($OrderPartnerOperationIn.Order_Input/ns0:state) = 0
or count($OrderPartnerOperationIn.Order_Input/ns0:city) = 0)
Depending on: OrderPartnerOperationIn
Located at:
\process\sequence[1]\if[1]\sequence[1]\if[1]\sequence[1]\if[1]
Condition 1: String representation: (true() =
$ValidateEMailOut.parameters/ns6:ValidateEMailResult)
Depending on: ValidateEMailOut
Located at: \process\sequence[1]\if[1]\sequence[1]\if[1]
Condition 2:
String representation:
($Validate_CreditCardOut.parameters/ns1:Validate_CreditCardResult = 1)
Depending on: Validate_CreditCardOut
Located at: \process\sequence[1]\if[1]
```

Listing 2: SOAP request

```
<SOAPenv:Envelope xsi:schemaLocation="..."
xmlns:xsi=http://www.w3.org/2001/XMLSchema-instance
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:SOAPenv="http://schemas.xmlSOAP.org/SOAP/envelope/"
xmlns:ord="..." xmlns:ord1="http://xml.netbeans.org/schema/OrderInfo">
<SOAPenv:Body>
  <ord:OrderPartnerOperation>
     <Order_Input>
       <ord1:firstname>?string?</ord1:firstname>
       <ord1:lastname>?string?</ord1:lastname>
       <ord1:email>?string?</ord1:email>
       <ord1:country>?string?</ord1:country>
       <!--Optional:-->
       <ord1:state>?string?</ord1:state>
       <!--Optional:-->
       <ord1:city>?string?</ord1:city>
       <ord1:postalCode>?string?</ord1:postalCode>
       <!--Optional:-->
       <ord1:address>?string?</ord1:address>
       <!--Optional:-->
       <ord1:phone>?string?</ord1:phone>
       <ord1:creditCardNumber>?string?</ord1:creditCardNumber>
       <ord1:total>?1.051732E7?</ord1:total>
       <ord1:currencyCode>?string?</ord1:currencyCode>
       <!--Optional:-->
       <ord1:errorCode>?string?</ord1:errorCode>
    </Order_Input>
  </ord:OrderPartnerOperation>
</SOAPenv:Body>
</SOAPenv:Envelope>
```

Listing 3: Input test data

```
<SOAPenv:Envelope xsi:schemaLocation="..."
xmlns:xsi=http://www.w3.org/2001/XMLSchema-instance
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:SOAPenv="http://schemas.xmlSOAP.org/SOAP/envelope/"
xmlns:ord="..." xmlns:ord1="http://xml.netbeans.org/schema/OrderInfo">
<SOAPenv:Body>
  <ord:OrderPartnerOperation>
     <Order_Input>
       <ord1:firstname>John</ord1: firstname>
       <ord1:lastname>Doe</ord1:lastname>
       <ord1:email>john@gmail.com</ord1:email>
       <ord1:country>USA</ord1:country>
       <ord1:postalCode>77590</ord1:postalCode>
       <!--Optional:-->
       <ord1:address>St. John 114</ord1:address>
       <!--Optional:-->
       <ord1:phone>00359111222</ord1:phone>
       <ord1:creditCardNumber>41111111</ord1:creditCardNumber>
       <ord1:total>40</ord1:total>
       <ord1:currencyCode>USD</ord1:currencyCode>
    </Order_Input>
  </ord:OrderPartnerOperation>
</SOAPenv:Body>
</SOAPenv:Envelope>
```

4) Generation of abstract test case for testing of business process's operation *OrderPartnerOperation*. The SOAP request to the partner web service is shown in Listing 2;

5) Creation of executable test cases from the abstract one by manual or automated specification of values for the input variables, considering the identified constraints on step 3. The expected results and assertions are defined using the test case generation tool. In order to execute the path that contains the activity *CityAssign*, the credit card number and the client email should be valid, while the city and the state may be omitted. The input data that satisfy these conditions as well as standard data types in the XSD schema of the business process are presented in Listing 3;

The following assertions are defined:

- HTTP status code – checks the HTTP status code of the response. The expected value is 200;
- XPath equals – the BPEL variable *ValidateEMailOut* should contain XPath expression *message/part/ValidateEMailResponse/ValidateEMailRe* with value *True*.
- XPath equals – the BPEL variable *Validate_CreditCardOut* should contain XPath expression *message/part/Validate_CreditCardResponse/Validate_CreditCardResult* with value *1*.
- Not XPath exists – the BPEL variable *OrderPartnerOperationIn* should not contain XPath expression *message/part/city*.
- Contains – the response is expected to contain the regular expression *<ord1:city>.*</ord1:city>*;
- Response time – the maximum value of the Response time is expected to be 5 000 ms.

6) The test cases are executed and the results are recorded and analyzed.

Table 1: Output from functional testing

No	Assertion	Result
1	HTTP status code	200
2	XPath equals on *ValidateEMailOut* variable	True
3	XPath equals on *Validate_CreditCardOut* variable	1
4	Not XPath exists on *OrderPartnerOperationIn* variable	XPath not found
5	Contains (case sensitive: "false", is regex: "true")	Value found
6	Response time	2562 ms

4.2. Isolation of partner web service

Since the partner web service *addresslookup* allows limited invocations over time, it is isolated from the business process by performing the following steps:

1) Selection of BPEL file, containing description of the *Check Payment* business process.

2) Selection of *addresslookup* web service for isolation;

3) Formal description of isolation, shown in Listing 4. The partner web service is executed though invoke activity *PostCodeInvoke*. The result from invocation is replaced with a string constant "TEXAS CITY".

4) Isolation of the partner service through transformation of the business process. Figure 2 shows the transformed part

of the business process, where the *Assign1* activity replaces the original *PostCodeInvoke* invoke activity.

Listing 4: Formal description of isolation

```
C: /CheckPayment/src/CheckPayment.bpel
---
process//invoke[@name='PostCodeInvoke']
$OrderPartnerOperationOut.Order_Output/ns0:city='TEXAS CITY'
$OrderPartnerOperationOut.Order_Output/ns0:state='TX'
---
```

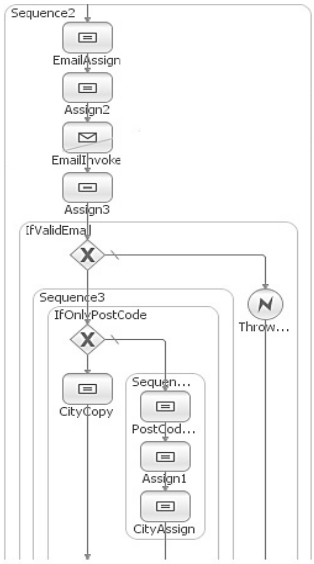


Figure 2: Isolation of partner web service

5) Execution of the transformed business process without invocation of *addresslookup* web service.

The output from isolation of partner web service is the same as one presented in Table 1. The partner web service *addresslookup* is successfully isolated without breaking of the normal execution of the business process.

4.3. Fault injection

The robustness testing of the business process is performed through the following steps:

1) Selection of BPEL file, containing description of the *Check Payment* business process.

2) Formal description of the fault that will be simulated, shown in Listing 5 – a delay in the communication channel between the business process and its partner web service *ValidatorDemo*.

Listing 5: Formal description of fault injection

```
C:/CheckPayment/src/CheckPayment.bpel
---
+proxy
wait=20
error_ratio=0
http://mathertel.de/AJAXEngine/S03 AJAXControls/ValidatorDemo.asmx
$ValidateEMailIn.parameters=$ValidateEMailOut.parameters
---
```

3) Transformation of the business process according to the formal description in step 3. The invocation of partner web service *ValidatorDemo* is replace with an invocation of proxy web service, called *PoxyInvoke*, which injects fault.

4) Execution of the transformed business process. The *ProxyInvoke* web service simulates a delay of 20 sec of the response from the partner web service.

The output from fault injection is the same as one presented in Table 1, except the value for the Response time. Due to delay of response from the partner web service *ValidatorDemo*, the actual Response time of the business process is 23718 ms.

5. Summary of methodology validation

TASSA methodology is evaluated based on the following assessment criteria:

- ***Applicability*** to SOA-based systems;
- ***Completeness*** of testing related to both WSCs and its components;
- ***Level of automation*** of the whole testing process;
- ***Level of novelty***.

More details about the results form evaluation of TASSA methodology can be found in [62,63,64,65]. They prove that the TASSA methodology provides solution of the *most significant problems* related to the WSC testing: (1) *Inability to instrument web services* that are not under control of the developer; (2) *Delay of testing* due to parallel development of the application components that have dependencies between each other; (3) *Lack automation of testing* for situations when the communication channel is interrupted or there is a noise in it and (4) *Difficulties during identification of the root causes* of failures and lack of a specific analysis approach dealing with different technology platforms.

The validation results shows that the isolation of partner web services can be applied to both synchronous and asynchronous WSCs. The isolation of the external dependencies and the injection of faults keep the original behavior of the WSC. In addition, the proposed approach for fault injection successfully simulate different types of failures. The identification and tracking the execution paths of WSC is supported by TASSA methodology through the approach for dependency analysis. This approach successfully identifies the conditional activities in the BPEL file and the variables, which values determine the execution on a given path.

The evaluation of TASSA methodology according to the above criteria proves its feasibility and effectiveness. The TASSA methodology proposes ***a novelty solution for testing SOA-based applications***. Its level of novelty is directly related to the maximum level of novelty of the adopted approaches. Therefore, the level of novelty of TASSA methodology is 3 from total 3. It delivers ***a complete solution***. Its completeness is calculated as a sum of completeness of the adopted approaches. The completeness of TASSA methodology is 11 from total 12. The TASSA methodology provides ***a high level of automation***. Its level of automation is computed as a sum of level of automation of the approaches that are proposed for its implementation. This metric varies due to possibility for combination of different approaches to achieve a specific testing goal. Since the testing approaches share test data, the manual activities related to its generation on the latest testing stages are minimized. The average level of automation of TASSA methodology is 7 from total 12. It is fully applicable to testing of WSCs, defined with WS-BPEL. The ***applicability*** of the TASSA methodology is directly related to the minimum applicability of the approaches that are proposed for its implementation. It depends on the concrete approaches used to test a given WSC. When all activities of TASSA methodology are

performed, the applicability is calculated as 1. When the activities related to isolation of partner web services and dependency analysis are not performed, the applicability is calculated as 3. In this case the TASSA methodology can be applied to a number of service-based applications.

Table 2 Summarizes the evaluation results. The first fifth rolls are related to the proposed testing approaches, while the last one shows the results for the TASSA methodology as a whole.

Table 2: Summary evaluation of TASSA methodology

Approach/ methodology	Novelty	Completeness	Automation	Applicability	Total
Isolation of dependencies	100%	100%	66%	25%	73%
Fault injection	66%	100%	66%	75%	77%
Dependencies analysis	100%	33-66%	33-66	25%	48-73%
Functional test case generation	33%	100%	100%	100%	66%
Test data generation	66%	66-100%	66%	75%	68-77%
Overall methodology	100%	80-93%	33-100%	25-100%	66-77%

6. Conclusion

The present work addresses the challenges of testing WSCs by proposing a novel methodology for testing, called TASSA. The methodology consists of the following main activities: (1) Checking for compliance with standards and isolation of partner web services; (2) Definition of testing goal and dependency analysis of the WSC to obtain testing paths; (3) Definition of functional test cases; (4) Definition of test cases for performance and load testing; (5) Definition of test cases for security testing; (6) Test case execution; and (7) Recording and reporting of test results. The optional execution of some activities enables focusing of the testing process on the specific user requirements and performing a step-by-step testing. For example, preparation activities could be performed at the beginning of business process development, including checking for compliance with standards. On the next step, when the partner web services are identified and are available, a functional testing could be performed. Finally, when the development of the business process under test is completed, non-functional testing could be conducted. A significant benefit of the TASSA methodology is that it deals with a small set of test artefacts – BPEL file, test data and test assertions. When the methodology is applied to the functional testing, a high level of automation can be achieved, in which only the BPEL file should be provided.

A lot of approaches are explored to identify the most suitable ones for implementation of the activities of TASSA methodology. New approaches are proposed for the most critical issues such as fault injection, isolation of partner services and dependency analysis. For other activities current approaches are adopted and are extended if needed. Two popular approaches for functional test case definition and test data generation are extended. The automation of TASSA methodology is shown onto a sample, yet realistic case study as a proof-of-concept. The complete validation is performed using different business processes in terms of number of activities, number of partner web services and type of communication – synchronous or asynchronous. The novelty, automation and applicability of TASSA methodology to testing serviced-based applications are evaluated regarding clearly defined assessment criteria. The results shows that it fully support end-to-end testing of WSCs covering all required testing activities.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research work has been supported by GATE “Big Data for Smart Society” project, funded by the European Union's Horizon 2020 WIDESPREAD-2018-2020 TEAMING Phase 2 programme under grant agreement no. 857155 and Big4Smart “Big Data Innovative Solutions for Smart Cities”, funded by the Bulgarian National Science fund, under agreement no. DN12/9 and agreement no. DN 02/11.

References

[1] Petrova-Antonova, S. Ilieva, D. Manova, “TASSA Methodology: End-to-end Testing of Web Service Compositions” in 11th International Conference on the Quality of Information and Communications Technology (QIATIC’2018), September 4-7, 2018, 264-267, Electronic ISBN: 978-1-5386-5841-3. https://doi.org/10.1109/QUATIC.2018.00046

[2] OASIS Web Services Business Process Execution Language, https://www.oasis-open.org/, last accessed 2019/11/04

[3] W3C., https://www.w3.org/standards, last accessed 2019/11/04

[4] T. Lertphumpanya, T. Senivongse, “Basis path test suite and testing process for WS-BPEL” WSEAS Transactions on Computers, 7(5), 483-496, 2008.

[5] Y. Yuan, Li, Z., W. Sun, “A graph-search based approach to BPEL4WS test generation” in IEEE Int. Conf. on Software Engineering Advances, 2006. https://doi.org/10.1109/ICSEA.2006.261270

[6] IBM - Rational Test Virtualization Server. (2016, January 1). http://www.ibm.com/software/products/en/rtvs, last accessed 2019/11/04

[7] Web Services API Mocking Overview | SOAP Mocking. (n.d.). https://www.soapui.org/soap-mocking/service-mocking-overview.html. last accessed 2019/11/04

[8] CA Service Virtualization - CA Technologies. (n.d.). https://www.ca.com/us/products/ca-service-virtualization.html, last accessed 2019/11/04

[9] R. Fletcher, “Betamax - Record & replay HTTP traffic”, http://betamax.software/, 2011, last accessed 2019/11/04

[10] Service Virtualization: Application & Data Simulation Software | Hewlett Packard Enterprise. (n.d.). http://www8.hp.com/us/en/software-solutions/service-virtualization/, last accessed 2019/11/04

[11] S. Bruning, S. Weissleder, M. Malek, “A fault taxonomy for service-oriented architecture” in IEEE 10th High Assurance Systems Engineering Symposium, 2007, 367-368. https://doi.org/10.1109/HASE.2007.46

[12] K. M. Chan, J. Bishop, J. Steyn, L. Baresi, S. Guinea, “A fault taxonomy for web service composition” in International Conference on Service-Oriented Computing, 2007, Lecture Notes in Computer Science, vol 4907. Springer, Berlin, Heidelberg, 363-375. https://doi.org/10.1007/978-3-540-93851-4_36

[13] C. Fu, B. G. Ryder, A. Milanova, D. Wonnacott, “Testing of java web services for robustness” ACM SIGSOFT Software Engineering Notes, 29(4), 2004, 23-34. https://doi.org/10.1145/1007512.1007516

[14] M. G. Fugini, B. Pernici, F. Ramoni, “Quality analysis of composed services through fault injection” Inf Syst Front (2009) 11: 227. https://doi.org/10.1007/s10796-008-9086-3.

[15] P. Kumar, Ratneshwer, “A Review on Dependency Analysis of SOA based System” in IEEE Fifth International Conference on Recent Trends in Information, Telecommunication and Computing, 2014, 69-81. https://doi.org/02.ITC.2014.5.11

[16] J. Yan, Z. Li, Y. Yuan, W. Sun, J. Zhang” Bpel4ws unit testing: Test case generation using a concurrent path analysis approach” In IEEE 17th International Symposium on Software Reliability Engineering, 2006, 75–84. https://doi.org/10.1109/ISSRE.2006.16

[17] Z. J. Li, H. F. Tan, H. H. Liu, J. Zhu, N. M. Mitsumori, “Business-process-driven gray-box SOA testing” IBM Systems Journal, 47(3), 2008, 457-472. https://doi.org/10.1147/sj.473.0457

[18] W. L. Dong, H. Yu, Y. B. Zhang, “Testing BPEL-based web service composition using high-level petri nets” in 10th IEEE International Enterprise Distributed Object Computing Conference, 2006. 441-444. https://doi.org/10.1109/EDOC.2006.59

[19] W. M. Van der Aalst, M. Dumas, C. Ouyang, A. Rozinat, E. Verbeek, "Conformance checking of service behavior" ACM Transactions on Internet Technology, 8(3), 13, 1-30, 2008. https://doi.org/10.1145/1361186.1361189
[20] M. Daghaghzadeh, M. Babamir, "An ABC based approach to test case generation for BPEL processes" in IEEE 3th International eConference on Computer and Knowledge Engineering, 2013, 272-277. https://doi.org/10.1109/ICCKE.2013.6682849
[21] H. Cao, S. Ying, D. Du "Towards model-based verification of BPEL with model checking" In Sixth IEEE International Conference on Computer and Information Technology, 2006, Seoul, 190-190. https://doi.org/10.1109/CIT.2006.185
[22] Z. Guangquan, R. Mei, Z. Jun, "A business process of web services testing method based on UML2. 0 activity diagram" in IEEE Workshop on Intelligent Information Technology Application, Zhang Jiajie, 2007, pp. 59-65. https://doi.org/10.1109/IITA.2007.83
[23] C. Bartolini, A Bertolino, E. Marchetti, A. Polini, "WS-TAXI: A WSDL-based testing tool for web services" In IEEE International Conference on Software Testing Verification and Validation, 2009, 326-335. https://doi.org/10.1109/ICST.2009.28
[24] C. Ma, C. Du, T. Zhang, F. Hu, X. Cai, "WSDL-based automated test data generation for web service" In International Conference on Computer Science and Software Engineering, 2008, 731-737. https://doi.org/10.1109/CSSE.2008.790
[25] H. M. Sneed, S. Huang "WSDLTest-a tool for testing web services" in Eighth IEEE International Symposium on Web Site Evolution, 2006, pp. 14-21. https://doi.org/10.1109/WSE.2006.24
[26] J. Offutt, W. Xu "Generating test cases for web services using data perturbation" ACM SIGSOFT Software Engineering Notes, 29(5), 2004, 1-https://doi.org/10. 10.1145/1022494.1022529
[27] X. Bai, W. Dong, W. T. Tsai, Y. Chen, "WSDL-based automatic test case generation for web services testing" in IEEE International Workshop on Service-Oriented System Engineering, 2005, 207-212. https://doi.org/10.1109/SOSE.2005.43
[28] DataGen - XML Test Data Generation Tool (2012). http://iwm.uni-koblenz.de/datagen/index.html, last accessed 2019/11/04
[29] Eclipse XML Editor Help - Eclipse Platform. (n.d.). http://help.eclipse.org/neon/index.jsp?topic=%2Forg.eclipse.wst.xmleditor.doc.user%2Ftopics%2Fcwxmledt.html, last accessed 2019/11/04
[30] M. Sorens (2009). Taking XML Validation to the Next Level: Introducing CAM. DevX. http://www.devx.com/xml/Article/41066, last accessed 2019/11/04
[31] V. Bergmann, Databene Benerator 0.7.6, manual, 2012.
[32] J. Durand, S. D. Green, S. Kulvatunyou, T. Rutt, "Test Assertions on steroids for XML artifacts" In Balisage: The Markup Conference, vol. 3, 2009.
[33] J. Tennison, "Validating XML with Schematron", Beginning XSLT, 2004, Apress.
[34] T. Bacon, J. Martin, "XMLUnit - Unit Testing XML for Java and .NET", http://www.xmlunit.org/, last accessed 2019/11/04
[35] Y. Zheng, J. Zhou, P. Krause "An Automatic Test Case Generation Framework for Web Services" Journal of Software, 2(3), 20017, 64-77. http://epubs.surrey.ac.uk/id/eprint/1975
[36] Ch. Ma, J. Wu, Tao Zh., Y. Zhang, X. Cai "Automatic Test Case Generation for BPEL Using Stream X-Machine" International Journal of u- and e-Service, Science and Technology, ISSN: 2005-4246, 2008, 27-36.
[37] Q. Yuan, J. Wu, Ch. Liu, L. Zhang "A model driven approach toward business process test case generation" in 10th International Symposium on Web Site Evolution, 2008. https://doi.org/10.1109/WSE.2008.4655394
[38] J. García-Fanjul, J. Tuya, Cl. de la Riva "Generating Test Cases Specifications for BPEL Compositions of Web Services Using SPIN" in International Workshop on Web Services Modelling and Testing, 2006, 83-94.
[39] W. Chen, J. Wei, G. Wu, X. Qiao' Developing a Concurrent Service Orchestration Engine Based on Event-Driven Architecture" Meersman R., Tari Z. (eds) On the Move to Meaningful Internet Systems: OTM 2008. OTM 2008. Lecture Notes in Computer Science, vol 5331. Springer, Berlin, Heidelberg, 2008, pp. 61-68.
[40] Apache ODE Team. Apache ODE BPEL Engine - Architectural Overview, 2013. http://ode.apache.org/developerguide/architectural-overview.html, last accessed 2019/11/04
[41] B. Li, X. Y. Tang, J. Lv, "The research and implementation of services discovery agent in web services composition framework". in IEEE International Conference on Machine Learning and Cybernetics, 2005, vol. 1, 78-84.
[42] S. Y. Hwang, H. Wang, J. Tang, J. Srivastava, "A probabilistic approach to modeling and estimating the QoS of web-services-based workflows", Information Sciences, 177(23), 2007, 5484-5503.
[43] N. Kabbani, S. Tilley, "Evaluating the capabilities of SOA security testing tools" In IEEE International Systems Conference, 2011, 129-134. https://doi.org/10.1109/SYSCON.2011.5929125
[44] R. Groenboom, R. Jaamour, "Securing Web Services" in OWASP Europe Conference, Leuven, Belgium, 2006
[45] OASIS Web Services Quality Factors Version 1.0. OASIS Committee Specification 01. http://docs.oasis-open.org/wsqm/WS-Quality-Factors/v1.0/cs01/WS-Quality-Factors-v1.0-cs01.html, last accessed 2019/11/04
[46] E. Yuan, J. Tong, "Attributed based access control (ABAC) for web services" In IEEE International Conference on Web Services, 2005. https://doi.org/10.1109/ICWS.2005.25
[47] M. Bartoletti P. Degano, G. L. Ferrari, "Plans for service composition" in Workshop on Issues in the Theory of Security, 2006.
[48] S. Haibo, H. Fan, "A context-aware role-based access control model for web services" in IEEE International Conference on e-Business Engineering, 2005, pp. 220-223. https://doi.org/10.1109/ICEBE.2005.1
[49] S. M. Babamir, F. S. Babamir S. Karimi, "Design and evaluation of a broker for secure web service composition" in IEEE International Symposium on Computer Networks and Distributed Systems, 2011, 222-226. https://doi.org/10.1109/CNDS.2011.5764577
[50] M. Anisetti, C. A. Ardagna, M. Bezzi, E. Damiani, S. P. Kaluvuri, A. Sabetta, "A Certification-Aware Service-Oriented Architecture" Bouguettaya A., Sheng Q., Daniel F. (eds) Advanced Web Services. Springer, New York, NY, 2014, 147-170.
[51] J. Biskup, B. Carminati, E. Ferrari, F. Muller, S. Wortmann, "Towards secure execution orders for composite web services" in IEEE International Conference on Web Services, 2007, 489-496.
[52] B. Zhou, D. Llewellyn-Jones, Q. Shi, M. Asim, M. Merabti, D. Lamb, "A Compose Language-Based Framework for Secure Service Composition" in IEEE International Conference on Cyber Security, 2012, 195-202. https://doi.org/10.1109/CyberSecurity.2012.32
[53] F. Martinelli, I. Matteucci, M. Petrocchi, L. Wiegand, "A formal support for collaborative data sharing", in International Conference on Availability, Reliability, and Security, Springer Berlin Heidelberg, 2012, 547-561.
[54] C. K. Yee, Design and Implementation of Test Case Generation Tool for BPEL Unit Testing (Doctoral dissertation, thesis), 2007.
[55] H. Roth, J. Schiefer, A. Schatten, "Probing and monitoring of WSBPEL processes with web services" in 3rd IEEE International Conference on E-Commerce Technology, 8th IEEE International Conference on and Enterprise Computing, E-Commerce, and E-Services, 2006, 30-30. https://doi.org/10.1109/CEC-EEE.2006.69
[56] L. Baresi, C. Ghezzi, S. Guinea, "Smart monitors for composed services" in ACM 2nd international conference on Service oriented computing, 2004, 193-202. https://doi.org/10.1145/1035167.1035195
[57] Active Endpoint, ActiveBPEL. www.active-endpoints.com/active-bpel-engineoverview, 2007. htm, last accessed 2019/11/04
[58] L. Jamen, SGhosh, "Oracle Fusion Middleware Performance and Tuning Guide 11g" https://docs.oracle.com/cd/E21764_01/core.1111/e10108/bpel.htm#ASPER99175, last accessed 2019/11/04
[59] C. Johnson, B. Newport, "Develop high performance J2EE threads with WebSphere Application Server" http://www.ibm.com/developerworks/websphere/techjournal/0606_johnson/0606_johnson.html, last accessed 2019/11/04
[60] M. Oriol Hilari, J. Marco Gómez, J. Franch Gutiérrez, D. Ameller, "Monitoring adaptable SOA systems using SALMon" in 1st Workshop on Monitoring, Adaptation and Beyond, 2008, 19-28.
[61] F. Barbon, P. Traverso, M. Pistore, M. Trainotti, "Run-time monitoring of instances and classes of web service compositions" in IEEE International Conference on Web Services, 2006, 63-71. https://doi.org/10.1109/ICWS.2006.113
[62] D. Manova, I. Manova, S. Ilieva, D. Petrova-Antonova, "faultInjector: A Tool for Injection of Faults in Synchronous WS-BPEL processes" in IEEE 2nd Eastern European Regional Conf. on the Engineering of Computer Based Systems, 2011, 99-105. https://doi.org/10.1109/ECBS-EERC.2011.23
[63] D. Petrova-Antonova, D. Manova, S. Ilieva, "TASSA: Testing Framework for web service orchestration". in IEEE/ACM 10th International Workshop on Automation of Software Test, 2015, 8-12. https://doi.org/10.1109/AST.2015.9
[64] D. Petrova-Antonova, S. Ilieva, D. Manova, "Automated Web Service Composition Testing as a Service" in Hammoudi S., Pires L., Selic B., Desfray P. (eds) Model-Driven Engineering and Software Development. MODELSWARD 2016. Communications in Computer and Information Science, vol 692. Springer, Cham, 114-131.
[65] I. Spassov, D. Petrova, V. Pavlov, S. Ilieva, "Data Dependency Analysis Tool for Web Service Business Processes" Murgante B., Gervasi O., Iglesias A., Taniar D., Apduhan B.O. (eds) Computational Science and Its Applications - ICCSA 2011. ICCSA 2011. Lecture Notes in Computer Science, vol 6786. Springer, Berlin, Heidelberg, 232-243.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 5, 163-173 (2019)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Chatting with Plants (Orchids) in Automated Smart Farming using IoT, Fuzzy Logic and Chatbot

Samruan Wiangsamut, Phatthanaphong Chomphuwiset, Suchart Khummanee*

Department of Computer Science, Faculty of Informatics, Mahasarakham University, 44150, Thailand

ARTICLE INFO

Article history:
Received: 16 August, 2019
Accepted: 25 September, 2019
Online: 10 October, 2019

Keywords:
Automated smart farming
Internet of Things (IoT)
Fuzzy logic
Chatbot

ABSTRACT

Plants are living organisms that can hear and recognize the environment around them but cannot communicate to inform their needs. Thus, in the past, humans thought that it was impossible to communicate with plants. However, in this modern era, humans can be able to communicate with these plants. In this paper, we propose a model that can interact (Chat) with plants cultivated in the automated farm system based on Internet of Things (IoT) and Fuzzy Logic. According to the communication of plants and humans, we apply a chatbot algorithm for sending/receiving messages between users and automated smart farming. The messages are processed by the natural language processing (NLP) to parse and interpret the meaning of the conversation. The experimented plant in this paper is orchid, namely Dendrobium Sonia (Bomjo). The result from the evaluation shows that the average accuracy (Harmonic mean) of chatting between the user and the orchid is equal to 0.71, the precision and recall are 0.75 and 0.6 respectively.

1 Introduction

What happens if we can talk to trees or plants? The answer is that we can know what they want, which will help us to respond to their needs accurately and precisely. For example, if the temperature rises, the moisture in the soil is decreased; as a result, plants need more water. If they are not watered at the right time, they will gradually wilt and eventually die. Besides, there is another exciting thing about plants, that is, the sound waves of different frequencies can affect the growth of different trees [1, 2]. Fahmi Huda et al. [3], for instance, discovered that acoustic waves could help stimulate the absorption of plant nutrients more efficiently, resulting in height, leaf area, stomatal opening, chlorophyll content, absorption of nitrogen, etc. of mustard Pakchoy are also higher.

The basic principle of chatting between plants and humans is applied and derived from the communication between ordinary and voiceless people. The deaf-mute can perceive various environments by looking, hearing, taste and touching, except for pronunciation, which is similar to the nature of plants. In fact, plants can recognize the various dangers that are approaching them, but they cannot response to aggressors. For example, the Mimosa pudica will shut their leaves if they are vibrated, or Morning glory flows will bloom or fold as the light intensity, etc. Therefore, we have adapted the principles mentioned above to apply to our chatting technique. Several researchers contributed methods for controlling and ordering various devices using the chat method via chatbots. Cyril Joe Baby et al. [4], for example, proposed a way to control various electrical devices in homes using a chatbot, working with IoT, that allow users to control devices within their house in a manual only. The system has not yet evaluated the errors regarding conversations via a chatbot and is not yet automated. Similar to Cyril's method, Muslih [5] introduced techniques for controlling electrical appliances in the workplace with IoT via Telegram Chatbot program in a smartphone. They claimed that the system can handle appliances correctly and also reduce the energy used, but they did not show the pros and cons that occurred from the communication between the IoT controller and the chatbot program. Besides, Angeline et al. [6] apply a chatbot implemented by Python to chat with customers when they need to buy goods in supermarkets, but it is only an extension, not an essential part of the system. Chatbots are applied to several ways [7], for example, Kaed et al. [8] developed an integrated system which collects data obtained from multiple IoT platforms running in the background, with a chatbot acting as a communication against the users in the foreground. Such a system is an intermediary that allows different IoT platforms to communicate with each other via a chatbot.

*Suchart Khummanee, Department of Computer Science, Faculty of Informatics, Mahasarakham University, Thailand. Email: suchart.k@msu.ac.th

In this paper, we propose techniques to chat with plants (Orchids) cultivated in the automated smart farming. The proposed techniques are based on IoT and Fuzzy logic using natural language processing (NLP) via the Python chatbot program. This paper is an extension of work initially presented in 2018 International Conference on Information Technology (InCIT) [9].

The rest of the paper is organized as follows: Section 2 depicts the background and related work. We articulate our methodology in section 3. The performance evaluation shows in the section 4. The system development is described in section 5. Section 6 shows our experimental results, and the last is the summary of this paper.

2 Background and Related Work

2.1 Stimulation and Recognition of Plants

Naturally, plants can recognize the stimulus from both inside and outside their bodies, which is a mechanism that occurs from the functioning of hormones or the various mechanisms of cells to make plants grow [10]. The stimuli influence the response of plants, which can be divided into two groups: external and internal stimuli. Plant's responses are different depending on these factors, such as the duration of stimulation of stimuli, the age and cell types that the sensory unit (Receptor), and the amount or concentrations of stimuli, etc. When plants are stimulated by stimuli, they can respond in several ways, such as movement, chemical release and tissue changes. For example, the leaves of Mimosa will be folded if there is a vibration near their leaves, as shown in Figure 1 (a). In this research, we are only interested in orchids because they are economic crops and have a high export value. Orchids (Dendrobium shown in Figure 1 (b)) can grow well if the natural stimulation factors: sunlight is 60 percent, the temperature is between 25 - 35 degrees Celsius, the relative humidity is between 60 - 70 percent and soil moisture is in the range of 60 - 70 percent [9].



(a) Mimosa (b) Orchid (Dendrobium)

Figure 1: Specimens of Mimosa and Orchid

2.2 Smart Farming with IoT

Smart farming is modern agriculture, which works with new technologies (IoT), robots, machines, etc. It has high accuracy for resolving problems of farmers by being friendly to the environment, being safe for consumers and using resources efficiently. For instance, Anukit and et al. [11] developed an alternative electric power supply from solar energy for a smart agricultural greenhouse, called the smart photovoltaic system, which is applied to the Lingzhi mushroom farm for saving energy costs. Besides, they [9, 12] used IoT to control the environment for the growth of orchids inside a smart farm due to unstable weather conditions, which allows farmers to plan production precisely. Hydroponics is another popular method today because farmers can get vegetables that are safe and easy to grow. Boonnam et al. [13] proposed a method of growing vegetables on a hydroponic farm controlled by IoT via mobile phones. The overall evaluation result from users' confidence in their system is higher than 70%. Next, farmers can make accurate decision making while undesirable events occur inside their farms based on real-time data using Agri-IoT [14] running on the semantic web technologies. IoT is not only restricted to use for crops, but it is also used for animal farms. For example, Memo et al. [15] adapted IoT to monitor various environments inside an animal farm, such the animal's waste, fire, water, animal's behaviours and others via a smartphone. The interesting researches related to IoT and the smart farming: [16, 17, 18].

2.3 Automatic Smart Farming with Fuzzy Logic

The automatic smart farm is a farm which can control various environments to suit each plant within the farm by itself without human intervention. For example, orchids grow well at a relative humidity (humid) ranging from 60 to 70% If humid around the farm is now 59%, therefore, the automatic smart farm must control the humid back to 60 to 70% by fogging automatically. To make any farm to be automated, there are a variety of methods, for a prevalent method, that is Fuzzy Logic (FL) [19]. Because of FL can explain ambiguous information so that humans can understand, especially the various environment information of plants. For example, three farmers choose unequal orchid growing temperatures, such as 25 - 30°C, 26 - 32°C and 25 - 32°C respectively. Therefore, these ambiguous data will be calculated using Fuzzy Logic. As a result, the appropriate temperature data range for orchid cultivation is 25 - 30°C. The output evaluated by the FL from multiple input levels is finite. There are four steps, as shown in Figure 2.

1. **Fuzzification** translates inputs, called crisp inputs into Fuzzy sets, and splits each crisp input into multiple levels, i.e., five levels = NH (Negative High), NL (Negative Low), Z (Zero), PL (Positive Low) and PH (Positive High) or three levels = L (Low), M (Medium) and H (High),

2. **Knowledgebase** creates rules that correspond to the problems being considered in the IF-THEN format, for example, IF $\text{humid}_{present}$=L AND humid_{Target}=M THEN Fogging,

3. **Inference engine** simulates human reasoning by using Fuzzy inference and knowledge base in conjunction with inputs,

4. **Defuzzification** transforms the Fuzzy set acquired from inference engine into a crisp output.

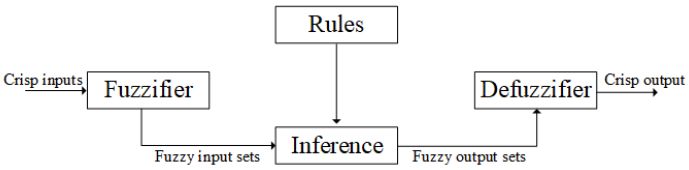


Figure 2: Fuzzy Logic workflow

2.4 Chatbots

A chatbot or virtual assistant is an artificial intelligence program [4, 7] which can simulate a conversation (Chat) with a user in natural language through messaging applications, websites or smartphones. Chatbot's structure consists of two parts: user request analysis and returning the response, as shown in Figure 3. Generally, chatbots work on three classification methods:

1. **Pattern Matching**: chatbots use a pattern matching method to group messages and create relevant responses from users. Artificial intelligence Markup language (AIML) is a standard structured model to create an interface which can be easily understood by humans response to users with the appropriate format, as shown in Figure 4.

2. **Natural Language Understanding (NLU)**: NLU is a subset of Natural Language Processing (NLP) which examines the group of texts or speeches (Sentences) and interprets the meaning specified. For example, a query for asking temperature inside Orchid smart farming on the 20th of July. NLU algorithm breaks down the query something like this: Orchid [intent] / need: temperature [intent] / inside farm [location] / July 20th [date].

3. **Natural Language Processing (NLP)**: NLP uses algorithms to analyze text so that the machine understands how humans speak. It is commonly used for text mining, machine translation, and automated question answering. NLP uses the machine learning algorithms which has the following steps: Tokenization, Sentiment Analysis, Normalization, Named Entity Recognition and Dependency Parsing, respectively [20].

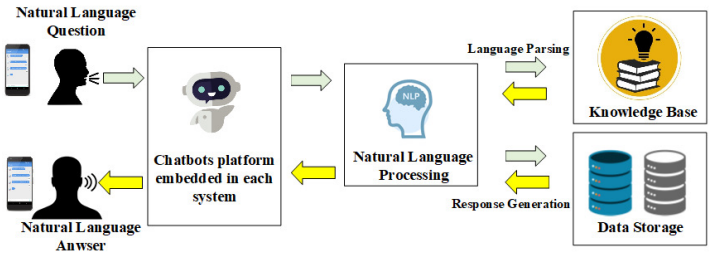


Figure 3: Chatbot's workflow

```
<aiml version="1.0" encoding ="UTF.8"?>
<category>
<pattern>HI HOW ARE YOU</pattern>
<template>
<srai>HI<srai><srai>HOW</srai><srai>ARE YOU</srai>
</template>
</category>
<category>
</aiml>
```

Figure 4: Artificial Intelligence Markup Language(AIML)

3 Proposed Methodology

Figure 5 illustrates the research procedures in order to build a chatbot for automated smart farming of Orchids. There are eight steps (1 to 8) as:

3.1 Collect information about orchid cultivation from farmers

The first step (1), we collect as much information as possible about the environments which are used to plant orchids from farmers including temperature (T), humidity (H) soil moisture (S) and sunlight (L). Information collected from orchid farmers in this research has eight sources listed in Table 1.

Table 1: Information on orchid cultivation from Orchid farmers

Farmer	Environment variables			
	Temperature (T)	Humidity (H)	Soil moisture (S)	Sunlight (L)
$\mathbf{F}_1$	25-30°C	N/A	N/A	N/A
$\mathbf{F}_2$	25-35°C	60-80%	60-80%	60-70%
$\mathbf{F}_3$	25-35°C	60-80%	60-80%	50-60%
$\mathbf{F}_4$	N/A	60-80%	N/A	N/A
$\mathbf{F}_5$	25-35°C	N/A	N/A	N/A
$\mathbf{F}_6$	25-35°C	50-70%	50-70%	N/A
$\mathbf{F}_7$	N/A	N/A	N/A	50-60%
$\mathbf{F}_8$	25-30°C	N/A	N/A	50-70%

Where F_i as any farmer, and N/A is unavailable.

According to Table 1, the first farmer (F_1) recorded only temperature (T) ranging from 25 to 30 degrees Celsius. Other information, the farmer does not record (N/A). In contrast, the farmer F_2 and F_3 recorded all the information of the environment variables: T ranges from 25 - 35 °C, H and S are between 60 - 80%, L is in the range of 60 - 70% and 50 - 60% respectively. Each type of information is calculated to determine membership values by the intersection operator ($\cap$), except the N/A as follows.

$$\begin{aligned} T_m &= T_{F_1} \cap T_{F_2} \cap T_{F_3} \cap T_{F_5} \cap T_{F_6} \cap T_{F_8} \\ &= [25, 30] \cap [25, 35] \cap [25, 35] \cap [25, 35] \cap \\ &\quad [25, 35] \cap [25, 30] = [25, 30] \end{aligned} \tag{1}$$

T_m is an intersection result of T_{F_i}, where i $\in$ [1, 8].

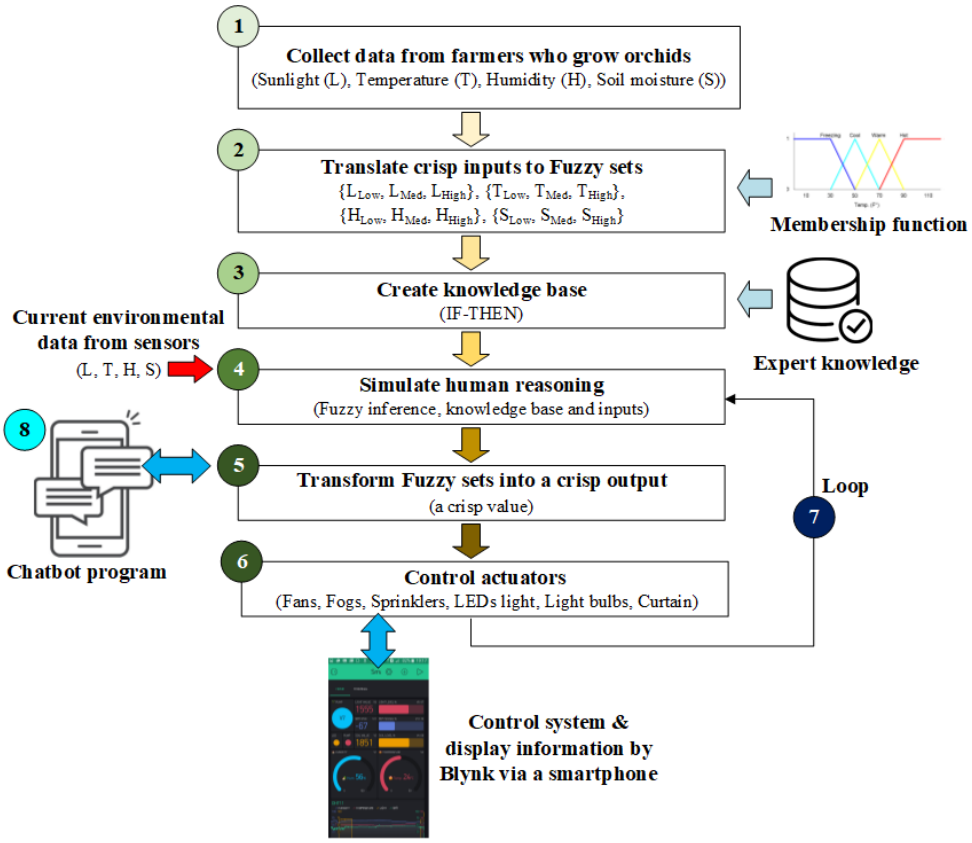


Figure 5: Block diagram for research methodology

$$H_m = H_{F_2} \cap H_{F_3} \cap H_{F_4} \cap H_{F_6} = [60, 80] \cap [60, 80] \cap [60, 80] \cap [50, 70] = [60, 70] \quad (2)$$

$$S_m = S_{F_2} \cap S_{F_3} \cap S_{F_6} = [60, 80] \cap [60, 80] \cap [50, 70] = [60, 70] \quad (3)$$

$$L_m = L_{F_2} \cap L_{F_3} \cap L_{F_7} \cap L_{F_8} = [60, 70] \cap [50, 60] \cap [50, 60] \cap [50, 70] = [60] \quad (4)$$

H_m, S_m and L_m are the intersection results of H_{F_i}, S_{F_i} and L_{F_i}, respectively.

3.2 Translate crisp inputs to Fuzzy sets

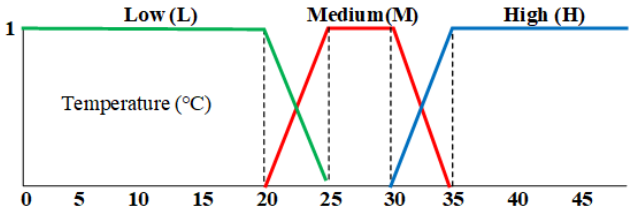
The second phase (2), crisp inputs are converted to Fuzzy sets, which are obtained from the first step. Theoretically, crisp inputs can be multiple levels, but in this research, only three levels are selected: T, H, S and L ∈ {LOW, MEDIUM, HIGH} because orchids are not very sensitive to climate changes. The membership functions are used to map crisp inputs to Fuzzy sets. They quantify linguistic term and demonstrate Fuzzy set graphically. A membership function for a Fuzzy set T on the universe of discussion X is defined as μT:X $\longrightarrow$ [0, 1]. Here, each element of X is mapped to a value between 0 and 1. Each membership function of T, H, S and L are illustrated in Figure 6 (a to d) respectively.

3.3 Create knowledge base rules

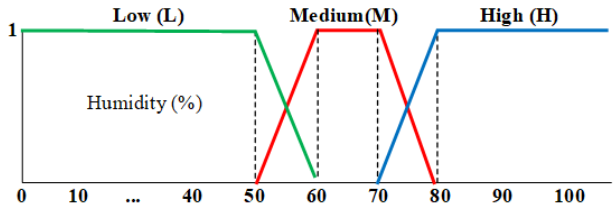
Next step, we map each membership function from the second step to a matrix, which is determined how we proceed if environment variables in smart farming occur in different situations. For example, from Figure 6 (a), if a current temperature ($T_{current}$) inside a smart farm is now equal to 24°C, but a target temperature (T_{target}) for planting orchids is in the middle (T ∈ [25, 30°C] = M); thus, the system needs to immediately reduce the current temperature to be the middle level (M) as shown in Table 2.

Another example, if the current relative humidity ($H_{current}$) is low about 58%, but the humidity suitable (H_{target}) for orchid planting is 60 - 80% (Middle). Therefore, the system will have to carry out fogging (Fog) to increase the humidity at an appropriate level immediately shown in Table 3. For other environment variables as S and L are mapped into a matrix, shown in Table 4 and 5 respectively.

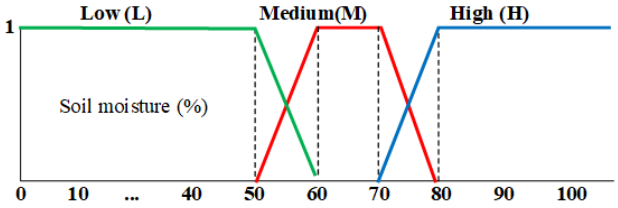
Next, we build a set of rules of T, H, S and L of each matrix in order to build the knowledge base in the form of IF-THEN structure like in Table 6. According to Table 6 rule No. 4, if the current humidity in the farm is high (H) and the orchids need moderate humidity (M), the system must turn on lamps to reduce the humidity inside the farm, and order the fans to suck moisture away from the farm simultaneously.



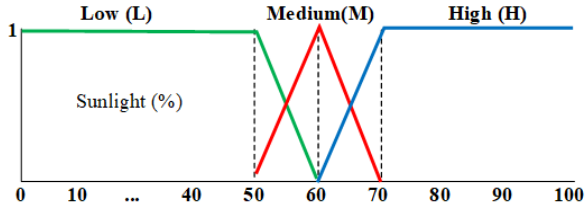
(a) Temperature (T)



(b) Humidity (H)



(c) Soil moisture (S)



(d) Sunlight (L)

Figure 6: Membership function of each crisp input

Table 2: The decision matrix of temperature

$\mathbf{T}_{target}$	$\mathbf{T}_{current}$		
	Low	**Medium**	**High**
Low	No change	Cool	Cool
Medium	Heat	No change	Cool
High	Heat	Heat	No change

T_{target} = target temperature for planting orchids, $T_{current}$ = current temperature inside a smart farming

Table 3: The decision matrix of humidity

$\mathbf{H}_{target}$	$\mathbf{H}_{current}$		
	Low	**Medium**	**High**
Low	No change	Heat/Fan	Heat/Fan
Medium	Fog	No change	Heat/Fan
High	Fog	Fog	No change

H_{target} = target humidity for planting orchids, $H_{current}$ = current humidity inside a smart farming

Table 4: The decision matrix of soil moisture

$\mathbf{S}_{target}$	$\mathbf{S}_{current}$		
	Low	**Medium**	**High**
Low	No change	No water	No water
Medium	Water	No change	No water
High	Water	Water	No change

S_{target} = target soil moisture for planting orchids, $S_{current}$ = current soil moisture in planting pots

Table 5: The decision matrix of sunlight

$\mathbf{L}_{target}$	$\mathbf{L}_{current}$		
	Low	**Medium**	**High**
Low	No change	Curtain	Curtain
Medium	Lighting	No change	Curtain
High	Lighting	Lighting	No change

L_{target} = target lighting for planting orchids, $L_{current}$ = current lighting inside a smart farming

3.4 Simulate human reasoning

Fuzzy set operations evaluate rules created in the previous step (Section 3.3). AND ($\wedge$) and OR ($\vee$) are used for describing a minimum (MIN) and a maximum value (MAX) respectively. Combine all results of the evaluation to form a final result, called a Fuzzy value. For example, the membership function of temperature shown in Figure 7 has three output Fuzzy sets: "Low", "Medium" and "High". Assuming that the temperature in the current orchid farm is equal to 23 °C, thus the membership value of the Fuzzy set "Low, Medium and High" are 0.35, 0.75 and 0, respectively. The Fuzzy set value is 0.35 $\vee$ 0.75 $\vee$ 0 = 0.75 (MAX). As a result, the Fuzzy value is Medium (M).

3.5 Transform the Fuzzy sets into a crisp output

Fuzzy set transformation to a crisp number is commonly called Defuzzification. It is a procedure that converts a Fuzzy set or Fuzzy number into a crisp value or number. The weighted average method (z^*) is the most popular used in fuzzy applications since it is one of the more computationally efficient methods and valid for symmetrical output membership functions. The algebraic expression gives z^*:

$$z^* = \frac{\sum_{i=1}^{n} \mu_{c_i}(x_i).(x_i)}{\sum_{i=1}^{n} \mu_{c_i}(x_i)} \quad (5)$$

Where c_1, c_2, ..., c_n are the output Fuzzy sets, and (x_i) is the value where the middle of each symmetric membership of the Fuzzy set c_i is observed.

Table 6: Fuzzy rule base of T, H, S and L

Rule No.	Condition	Action
1	**IF** $T_{current}$ = {L} **AND** T_{target} = {M} **THEN**	Heat
2	**IF** $T_{current}$ = {H} **AND** T_{target} = {M} **THEN**	Cool
3	**IF** $H_{current}$ = {L} **AND** H_{target} = {M} **THEN**	Fog
4	**IF** $H_{current}$ = {H} **AND** H_{target} = {M} **THEN**	Heat/Fan
5	**IF** $S_{current}$ = {L} **AND** S_{target} = {M} **THEN**	Water
6	**IF** $S_{current}$ = {H} **AND** S_{target} = {M} **THEN**	No water
7	**IF** $L_{current}$ = {L} **AND** L_{target} = {M} **THEN**	Lighting
8	**IF** $L_{current}$ = {H} **AND** L_{target} = {M} **THEN**	Curtain

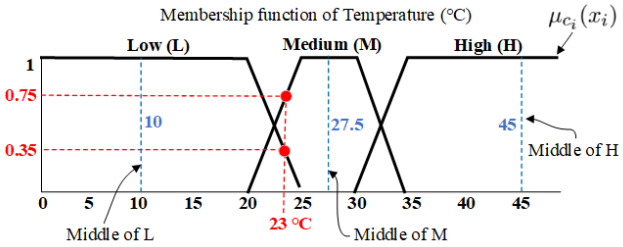


Figure 7: Fuzzy defuzzification

In Figure 7, if $\mu_{c_i}(x_i) = 1$, $x_{Low} = 10°C$, $x_{Medium} = 27.5°C$ and $x_{High} = 45°C$, we can estimate a crisp output of the temperature by using the z^* as follows:

$$z^* = \frac{(1).(10) + (1).(27.5) + (1).(45)}{(1+1+1))} = 27.5°C$$

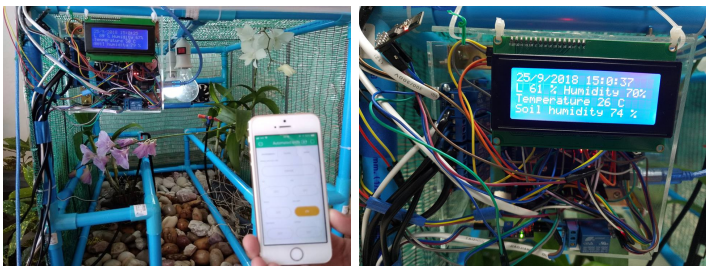
Also, the crisp output of humidity, soil moisture and sunlight: 65%, 65% and 60% respectively.

3.6 Control actuators

After the defuzzification process is executed successfully, the Fuzzy set value and crisp output will be used to control the actuators inside an orchid farm. For example, if the environment variables inside the orchid farm are now: $T_{current}$ = 21 °C (L = Low), $H_{current}$ = 54% (L), $S_{current}$ = 54% (L) and $L_{current}$ = 66% (H = High), the equipment (Actuators) inside the orchid farm must carry out the following tasks: Light bulbs = on (Heat), Fogs = on (Water), Sprinkles = on (Water) and Curtain = on (Sheltered greenhouse). The overview of the demonstration farm for orchid cultivation is shown in Figure 8. The methods of controlling the system can be done in two ways: allowing the system to operate automatically or manually via the Blynk software [21] on the mobile phone.

3.7 Loop control

The system is designed to work automatically, based on the knowledge that has been created from the previous steps, but users can still control the system by manual handling. The system reads various environmental data from the sensors in the fourth step and then uses such data to evaluate a Fuzzy value in the fourth step and a crisp value in the fifth step. The Fuzzy value will be calculated with the appropriate knowledge based rules (IF-THEN rules from the third step) for orchid planting. After that, the result will be taken to control actuators inside the smart orchid farming. The system works in a continuous loop without interruption.



(a) Controlling the system via a smartphone (b) Actuators inside the demonstration farm

Figure 8: Overview of demo orchid farming

3.8 Chat to Orchids (OrchidChatbot)

At the last phase (8), we design and create a chatbot to be used for talking to orchids to inquire about how healthy orchids are. The overview of chatbot's work is shown in Figure 9. According to Figure 9 (A), we use ChatterBot [22] running on Python version 3.7 to chat with orchids, which are grown inside the farm (D), through Blynk servers on the cloud (B), which act as intermediaries for communication. Communication between our chatbot (OrchidChatbot) and the Blynk servers to retrieve the environment information of the orchid farm, we use the Blynk library for Python [23]. OrchidChatbot starting from the smart farm control system (D) sends a Fuzzy value of the temperature (T), humidity (H), soil the moisture (S) and sunlight (L) to store in the Blynk servers by using Blynk libraries (C), these Fuzzy values must be evaluated by Fuzzy Logic already (C). Next, OrchidChatbot retrieves environment information (T, H, S and L) from Blynk serves for processing natural language like Figure 3 together with the Orchid corpus as shown in Figure 10. The format of the ChatterBot corpus explains in Table 7, and OrchidChatbot program, which is coded in Python, is shown in algorithm 1.

Referring to the chatbot program in algorithm 1, the first line loads Blynk libraries to a chatbot for communicating against Blynk servers on the cloud. Line number 2 is importing the Chatbot module from the chatterbot library to

build our chatbot program. The user authentication to Blynk servers by a key is shown in the line number 3. Creating virtual pins to send/receive the temperature, humidity, soil moisture and light between our chatbot and the Blynk servers shows in the line 4 to 7 respectively. Next, the line number 8 and 9 creates a chatbot and trainer object; and the line 10 points to the address of the Orchid corpus for training data. The chatbot retrieves the environment values from the Blynk servers from the line number 13 to 16. Last, line number 17 to 30 are used to respond between users and the chatbot; and when users want to leave the chat, type "bye" (line number 31 - 33).

```
categories:
- health
conversations:
- - How are you?[T:L]
  - I am feeling cold
- - How are you?[T:M]
  - I am feling warm.
- - How are you?[T:H]
  - I'm feeling hot.
- - How are you?[H:L]
  - I am feeling dry.
- - How are you?[H:M]
  - I am feeling cool, being good.
- - How are you?[H:H]
  - I am feeling wet.
- - How are you?[S:L]
  - I am feeling that my roots are dry.
- - How are you?[S:M]
  - I feel that my roots are fine.
- - How are you?[S:H]
  - I feel that my roots are too wet.
- - Can I help you?[T:L]
  - Yes, I want the warm weather.
- - Can I help you?[T:M]
  - No, it is already warm.
- - Can I help you?[T:H]
  - Yes, I want you to fog to reduce my heat.
- - Can I help you?[H:L]
  - Yes, I want more heat.
- - Can I help you?[H:M]
```

Figure 10: Examples of Orchid corpus

4 Performance evaluation

The performance evaluation in this research is divided into three components: 1) growth rate of orchid inflorescence length (cm./a week) planing inside the smart farming for ten weeks; 2) satisfaction from users who control and view the information via smartphones by the average ($\bar{x}$); and the accuracy of the conversation between the users, chatbot and the orchids by Dialogue specification and Harmonic mean, which the equations are shown as follows:

The average ($\bar{x}$):

$$\bar{x} = \frac{1}{n}.\sum_{i=1}^{n} x_i \tag{6}$$

Where $x_1, x_2, ..., x_n$ are the observed values of the sample testers, $\bar{x}$ is the mean value of these observations, and n is the number of observations in the sample. In this research, the number of testers (n) is thirty.

Algorithm 1 OrchidChatbot

```
1: import Blynk libraries
2: import ChatBot libraries
3: SET auth_token by 'AuthToken'
4: SET Temperature by Blynk_virtual_pin(1)
5: SET Humidity by Blynk_virtual_pin(2)
6: SET Soil_moisture by Blynk_virtual_pin(3)
7: SET Sunlight by Blynk_virtual_pin(4)
8: Create a chatbot object from ChatBot module
9: Create a trainer object from ChatterBotCorpusTrainer
10: SET the Orchid corpus path directory (namely orchid.yml)
11: Training Orchid corpus
12: while True:
13:     Temp ⟵ get temperature from Blynk server
14:     Humid ⟵ get Humidity from Blynk server
15:     Soil ⟵ get Soil_moisture from Blynk server
16:     Light ⟵ get Sunlight from Blynk server
17:     message = input('you:')
18:     if message.strip()!='bye':
19:         if Temp != NULL:
20:             reply = chatbot.get_response(message+Temp)
21:             print('ChatBot :', reply)
22:         if Humid != NULL:
23:             reply = chatbot.get_response(message+Humid)
24:             print('ChatBot :', reply)
25:         if Soil != NULL:
26:             reply = chatbot.get_response(message+Soil)
27:             print('ChatBot :', reply)
28:         if Light != NULL:
29:             reply = chatbot.get_response(message+Light)
30:             print('ChatBot :', reply)
31:     if message.strip()=='bye':
32:         print('ChatBot: bye')
33:         break
```

The Dialogue specification:

$$Dialogue\ specification = \frac{\sum_{j=1}^{m} wc_j}{\sum_{i=1}^{n} wa_i}; n >= m \tag{7}$$

Where wc_j are the words that chatbot answered correctly, and wa_i are all the words that chatbot answered.

The harmonic mean:

$$H = \left(\frac{\sum_{i=1}^{n} x_i^{-1}}{n}\right)^{-1} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} \tag{8}$$

Where $x_1, x_2, ..., x_n$ are positive real numbers. For example, suppose that the user and chatbot chat together, as shown in Table 8, the results from the calculation with equation 7 and 8:

Table 8: Examples of conversation between the user and chatbot

User asked	Chatbot answered	Possible answers in Corpus
How are you?	I am cold.	I am feeling cold. I am feeling warm. I'm feeling hot.
Are you hot?	Yes, I am hot	No, I am so happy. Yes, I am very hot. No, I am so so.

The dialogue specification of question 1 and 2 are equal to 1.0 because the chatbot can answer both questions correctly. The harmonic mean of both questions:

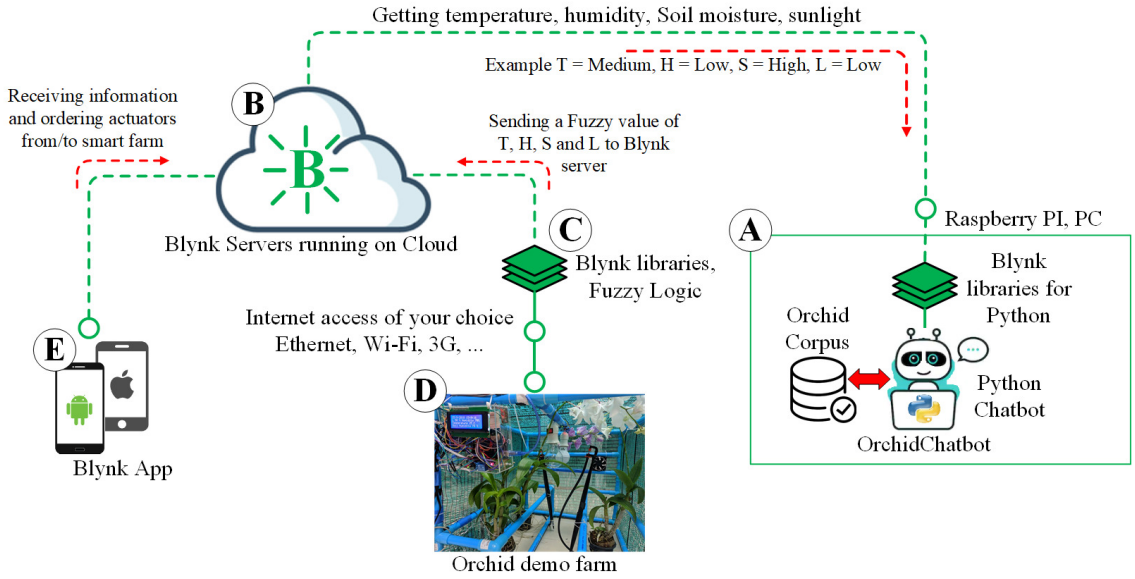


Figure 9: The working process of chatbot in Orchid demo farm

Table 7: The format of ChatterBot corpus

Syntax	Description	Example
categories:	Category of questions	categories: - health
conversations:	Start the conversation of each question category	conversations: - - How is your health? - I'm fine.
- - the question[?]	Questions that users ask to Chatbot	- - Hello - - How are you?
- the answer[.]	The answers which the Chatbot replies to the users	- Hi - I'm also good.
[X:Y]	Fuzzy value of each environment variable	- [T:L] - [H:M]

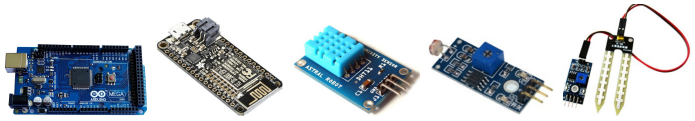
$$H = \frac{2}{\sum_{i=1}^{2} \frac{1}{x_i}} = \frac{2}{\frac{1}{x_1} + \frac{1}{x_2}} = \frac{2}{\frac{1}{1} + \frac{1}{1}} = 1.0$$

We also use the *precision* and *recall* methods, which are the classic techniques to evaluate a chatbot chatting. We create a total of 60 questions (Orchid corpus) in the discussion between users and orchids in this research listed in Table 10.

5 System development

5.1 Hardware for smart orchid farming

The devices such as sensors and actuators used in the orchid farm are as follows. Arduino Mega 2560 and Esp8266 module is used to control the entire system of automated smart farming. According to the sensors: the temperature and humidity (DHT11), soil moisture and LDR photo-resistor sensor as illustrated in Figure 11. The actuators are water pumps (12V DC), cooling fan (220V 50Hz), light bulbs (220V 50W), motors (12V DC), LED bulbs (220V 400 lumen), fogs and sprinkles as illustrated in Figure 12.



(a) Maga2560 (b) esp8266 (c) DHT11 (d) LDR (e) Soil

Figure 11: Arduino Mega, ESP8266 and sensors



(a) Light/LED (b) Fan (c) Pump (d) Sprinkler (e) Fog

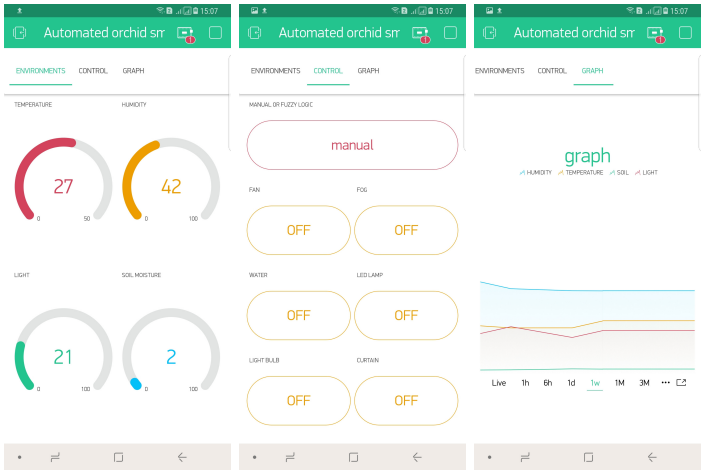
Figure 12: Actuators working in smart farm

5.2 Software for smart orchid farming

Orchid smart farming is controlled by a software developed from Blynk [21] application. It can be configured in two modes: the automatic and manual mode. According to the manual mode, orchid farmers can command the system via a smartphone by themselves. Another mode can operate automatically by itself. The sample interfaces of Blynk represent in Figure 13. For example, Figure 13 (a) displays the information in the Guage style, the operation buttons of various devices are shown in Figure 13(b) and Figure 13(c) shows usage statistics in the graph style.

5.3 OrchidChatBot hardware and software

The hardware and software on the chatbot side consist of the Linux operating system running on the Raspberry PI, Python 3.7, Blynk libraries for Python and ChatterBot for Python as illustrated in Figure 9 (A). The OrchidChatBot algorithm is explained in Algorithm 1.



(a) Gauge reports (b) Manual control (c) Graph reports

Figure 13: Blynk software for controlling the orchid farming

6 The experimental results

The results of the system evaluation are divided into three parts:

6.1 Results of system control via a mobile

Controlling smart orchid farming via mobile phones has been tested against 30 users. The results from the following tests shown in Table 9. Overall, the control system developed from Blynk has an average above 4.5 almost every item because the interfaces are developed to be exciting and easy to use.

Table 9: Satisfaction evaluation of users

Topic of evaluation	Satisfaction value ($\bar{x}$)
System design satisfaction	
1. Interesting system	4.8
2. Modernization of the system	4.7
3. Easy to use and easy to understand	4.9
4. System simulation is appropriate	4.4
Satisfaction in using the system	
1. Accuracy of the system	4.7
2. Appropriate to needs of users	4.6

6.2 Results of orchids planting inside the smart farming

The orchid farm is tested for a total of eight weeks from April to May 2018. We evaluate the growth of the inflorescence length of the orchids inside the farm every day. At night, the lighting system stops working to allow the orchids to get rid of the carbon dioxide. The growth statistics obtained from the experiments are shown in Figure 14. According to the line chart, the growth rate of the inflorescence length under a traditional farm is not stable because orchids proliferate

during the rainy season. The length of the inflorescence increases by an average of 34.3 centimeters (cm.) within ten weeks. In contrast, orchids are grown under an automatic smart farm; the growth rate is relatively constant at about 27.38 cm. only. Although in the overall, the growth rate of orchids inflorescence length raised in a traditional farm is higher than automated smart farm, the steady growth rate inside an automated smart farming can control the yield of orchids better than in practice. Since farmers can evaluate the production of orchids more precisely.

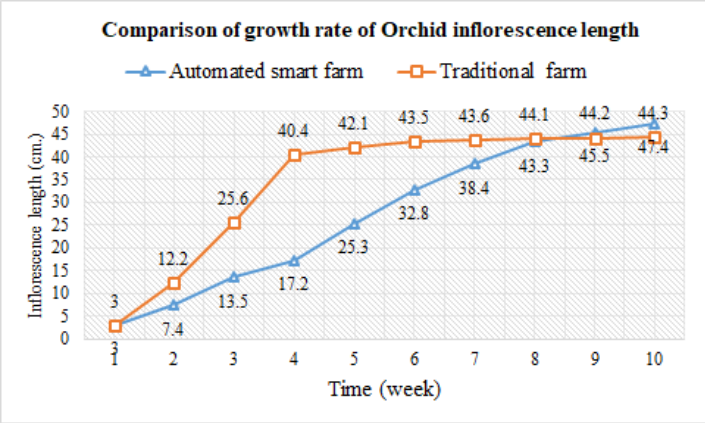


Figure 14: Comparison of the growth rate of the orchids inside the traditional farm against an automated smart farming

6.3 *Results of chatting between orchids and our chatbot*

The conversation between the orchids and chatbot is carried out through the command-line interface of Python shell, shown in Figure 15. Messages used to chat are related to the health of the orchids based on the information that the sensors can detect. Examples of questions and possible answers used in this research listed in Table 10.

The conversation is evaluated using the dialogue specification [24] and harmonic mean in the equation (7, 8). The average accuracy (Harmonic mean) of chatting between the user and the orchid is equal to 0.71, the *precision* = 0.75 and *recall* = 0.6 respectively.

7 Conclusion and future work

In this paper, we propose a chatbot which is used to chat with orchids planting inside the automated smart farming controlled by IoT and Fuzzy Logic, namely OrchidChatBot. The results of the system testing are divided into three parts: the average satisfaction of users using the mobile application to control an automated smart farming is equal to 4.68, the average growth rate of orchid inflorescence cultivated in the smart farm is 27.38 cm., and the average accuracy of the conversation between the chatbot and orchids are 0.71 (harmonic mean), 0.75 (precision) and 0.6 (recall) respectively. Besides, the significant advantages of automatic smart farming are reducing fertilizer, water, chemicals, insects, and yield control.

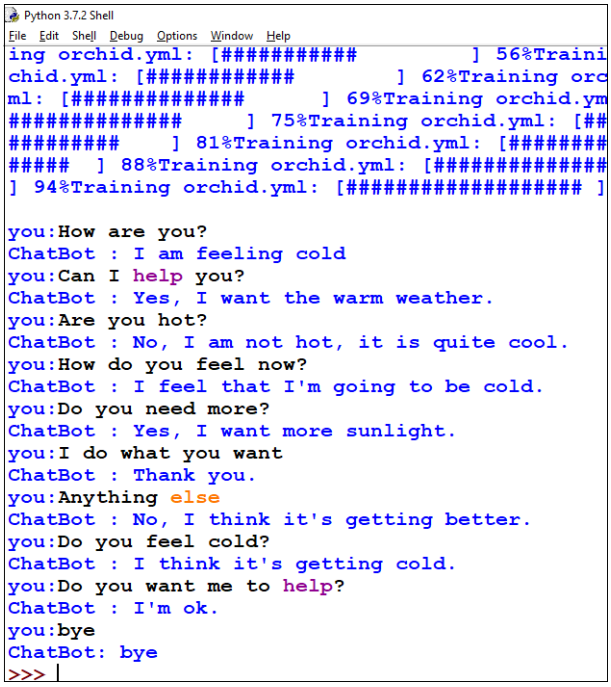


Figure 15: Chatting between Orchids and the chatbot by the command-line interface

Table 10: Examples of questions and possible answers

Question	Possible answers
How are you?	I am feeling cold. I am feeling warm. I'm feeling hot. I am feeling dry. I am feeling cool, being good. I am feeling wet. I am feeling that my roots are dry. I feel that my roots are fine. I feel that my roots are too wet.
Can I help you?	Yes, I want the warm weather. No, it is already warm. Yes, I want you to fog to reduce my heat. Yes, I want more heat. No, I am already invigorating. Yes, I want a blanket to increase warmth.
How do you feel now?	I feel that I'm going to be cold. I feel that I'm very comfortable. I feel that the weather is getting hotter.
Do you need more?	Yes, I want more sunlight. No, I do not more. Yes, I want to reduce the sunshine.
...	...
Are you hot?	No, I am so happy. Yes, I am very hot. No, I am not hot, it is quite cool. No, I am so so.

Moreover, chatting with orchids will help farmers to recognize the health of orchids, and to resolve problems that will occur with orchids more accurately.

In the future, we will enhance the conversation between the chatbot and orchids by using voice instead the command-line interface.

References

[1] Anindita Roy Chowdhury and Anshu Gupta, "Effect of Music on Plants – An Overview" International Journal of Integrative Sciences,

Innovation and Technology (IJIIT)., **4**(6), 30 – 34, 2015.

[2] Munasinghe, D. S. P, Weerakoon, S. R, Somaratne, S, "The effect of Buddhist pirith chanting and Western pop music on growth performance of "Pranajeewa", Codariocalyx motorius (Houtt.) H. Ohashi" Ceylon Journal of Science, 47(4), 357–361 2018, DOI: http://doi.org/10.4038/cjs.v47i4.7553

[3] F. H. Zakariya, M. Rivai and N. Aini, "Effect of automatic plant acoustic frequency technology (PAFT) on mustard pakcoy (Brassica rapa var. parachinensis) plant using temperature and humidity parameters," 2017 International Seminar on Intelligent Technology and Its Applications (ISITIA), Surabaya, 2017, pp. 334-339. doi: 10.1109/ISITIA.2017.8124105

[4] C. J. Baby, F. A. Khan and J. N. Swathi, "Home automation using IoT and a chatbot using natural language processing," 2017 Innovations in Power and Advanced Computing Technologies (i-PACT), Vellore, 2017, pp. 1-6. doi: 10.1109/IPACT.2017.8245185

[5] M. Muslih et al., "Developing Smart Workspace Based IOT with Artificial Intelligence Using Telegram Chatbot," 2018 International Conference on Computing, Engineering, and Design (ICCED), Bangkok, Thailand, 2018, pp. 230-234. doi: 10.1109/ICCED.2018.00052

[6] R. Angeline, T. Gaurav, P. Rampuriya and S. Dey, "Supermarket Automation with Chatbot and Face Recognition Using IoT and AI," 2018 3rd International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2018, pp. 1183-1186. doi: 10.1109/CESYS.2018.8723978

[7] B. Kohli, T. Choudhury, S. Sharma and P. Kumar, "A Platform for Human-Chatbot Interaction Using Python," 2018 Second International Conference on Green Computing and Internet of Things (ICGCIoT), Bangalore, India, 2018, pp. 439-444. doi: 10.1109/ICGCIoT.2018.8753031

[8] C. E. Kaed, A. Ponnouradjane and D. Shah, "A Semantic Based Multi-Platform IoT Integration Approach from Sensors to Chatbots," 2018 Global Internet of Things Summit (GIoTS), Bilbao, 2018, pp. 1-6. doi: 10.1109/GIOTS.2018.8534520

[9] S. Khummanee, S. Wiangsamut, P. Sorntepa and C. Jaiboon, "Automated Smart Farming for Orchids with the Internet of Things and Fuzzy Logic," 2018 International Conference on Information Technology (InCIT), Khon Kaen, 2018, pp. 1-6. doi: 10.23919/INCIT.2018.8584881

[10] Chamovitz, D., What a Plant Knows: A Field Guide to the Senses, Scientific American, 2013.

[11] A. Saokaew, O. Chieochan and E. Boonchieng, "A smart photovoltaic system with Internet of Thing: A case study of the smart agricultural greenhouse," 2018 10th International Conference on Knowledge and Smart Technology (KST), Chiang Mai, 2018, pp. 225-230. doi: 10.1109/KST.2018.8426071

[12] F. A. Rafi, A. Kowanda, N. S. Salahuddin and T. Saptriani, "Design of Orchid Monitoring System Based on IoT," 2018 Third International Conference on Informatics and Computing (ICIC), Palembang, Indonesia, 2018, pp. 1-6. doi: 10.1109/IAC.2018.8780419

[13] N. Boonnam, J. Pitakphongmetha, S. Kajornkasirat, T. Horanont, D. Somkiadcharoen, J. Prapakornpilai, "Optimal Plant Growth in Smart Farm Hydroponics System using the Integration of Wireless Sensor Networks into Internet of Things", Advances in Science, Technology and Engineering Systems Journal, vol. 2, no. 3, pp. 1006-1012 (2017).

[14] A. Kamilaris, F. Gao, F. X. Prenafeta-Boldu and M. I. Ali, "Agri-IoT: A semantic framework for Internet of Things-enabled smart farming applications," 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT), Reston, VA, 2016, pp. 442-447. doi: 10.1109/WF-IoT.2016.7845467

[15] M. H. Memon, W. Kumar, A. Memon, B. S. Chowdhry, M. Aamir and P. Kumar, "Internet of Things (IoT) enabled smart animal farm," 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2016, pp. 2067-2072.

[16] R. C. Andrew, R. Malekian and D. C. Bogatinoska, "IoT solutions for precision agriculture," 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, 2018, pp. 0345-0349. doi: 10.23919/MIPRO.2018.8400066

[17] N. Putjaika, S. Phusae, A. Chen-Im, P. Phunchongharn and K. Akkarajitsakul, "A control system in an intelligent farming by using arduino technology," 2016 Fifth ICT International Student Project Conference (ICT-ISPC), Nakhon Pathom, 2016, pp. 53-56. doi: 10.1109/ICT-ISPC.2016.7519234

[18] S. Suakanto, V. J. L. Engel, M. Hutagalung and D. Angela, "Sensor networks data acquisition and task management for decision support of smart farming," 2016 International Conference on Information Technology Systems and Innovation (ICITSI), Bandung, 2016, pp. 1-5. doi: 10.1109/ICITSI.2016.7858233

[19] Melin P., Castillo O., Kacprzyk J., Reformat M., Melek W., Fuzzy Logic in Intelligent System Design, Springer International Publishing, 2018.

[20] Sudhir K. Mishra, Artificial Intelligence and Natural Language Processing, Cambridge Scholars Publisher, 2018.

[21] Blynk. (2018, 12 Oct). The mobile app for the IOT. [Online]. Available: https://www.blynk.cc/

[22] ChatterBot. (2019, 10 Aug). How ChatterBot Works. [Online]. Available: https://chatterbot.readthedocs.io/en/stable/

[23] Python client for Blynk. (2019, 10 Aug). Blynk library for Python 2, Python 3, and MicroPython. [Online]. Available: https://github.com/vshymanskyy/blynk-library-python

[24] S. Ghose and J. J. Barua, "Toward the implementation of a topic specific dialogue based natural language chatbot as an undergraduate advisor," 2013 International Conference on Informatics, Electronics and Vision (ICIEV), Dhaka, 2013, pp. 1-5. doi: 10.1109/ICIEV.2013.6572650

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 169-173 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Study on Development of Information and Communication Ethics Sensitivity Measurement for Elementary School Students

Woochun Jun*

Seoul National University of Education, Department of Computer Education, 06639, Korea

ARTICLE INFO

Article history:
Received: 17 November, 2019
Accepted: 22 December, 2019
Online: 22 January, 2020

Keywords:
ICT Ethics
Sensitivity
Elementary School Students
ICT Ethics
Software Education

ABSTRACT

With advances in information and communication technologies and internet technologies, we have many benefits in our daily life. Most people have smart devices and use them for their work and leisure, and so on. However, these advances technologies brought us various side effects such as copyright violation, internet addiction, cybercrimes, personal information infringement, etc. In order to cope with those side effects, it is very important to teach information and communication ethics to students as early as possible. In addition, in order to teach young students, first, their awareness to information and communication ethics needs to be tested. In the literature, most awareness test tools are to test knowledge of information and communication ethics. In fact, those awareness tools have not consider sensitivity for information and communication ethics. In this work, an extensive and integrated sensitivity measurement tool is developed. At first sensitivity measurement standards are developed based on literature review works. Many elementary school teachers and ICT education experts survey the initial standards. Finally, sensitivity measurement standards are finalized with 6 areas and 19 standards. It is hoped that the proposed standards will be very helpful to diagnose awareness of elementary school students for information and communication ethics.

1. Introduction

In current knowledge and information society, everybody can have benefits from advances in information and communication technologies(ICT), internet technologies, and smart technologies, etc. Those advances technologies have brought many benefits in our daily life. Now many people have their own personal and smart device such as tablet PC, and can access to news and stock prices as quickly as possible.

Education area is also changing very quickly. In current classrooms, many PC and mobile devices are used for students and teachers. Students can get information anytime anywhere and exchange their ideas with anybody in the world. Recently, with advanced VR(Virtual Reality) and AR(Augmented Reality) technologies, VR/AR-based education becomes popular[1,2,3].

As we have more benefits from advanced technologies, side effects form those technologies become serious. The typical side effects include personal information infringement, copyright violation, cybercrimes, hacking, and so on. Most countries provide precaution and cure plans for those who suffering side effects. Usually costs for cure are tremendous. Thus, it is necessary to prevent those side effects as early as possible. The fundamental step to prevent is through education. Many countries provide education in the form of information and communication ethics education.

Recently robots are used in many areas as artificial intelligence and machine learning technologies have been developed. In this sense, ICT ethics education is extended to robot ethics education[4,5,6]. Robot ethics education can have big interest in the future as robots will be used more in industry and home. We need more fine and detailed principles and codes of ethics for more advanced robot.

In order to prevent and cure side effects from advanced technologies, it is very important to diagnose the status and awareness of students for ICT ethics. Based on correct diagnosis, it is possible to provide appropriate plans for precaution and cure. In the literature, awareness measurements are based on sole knowledge level of students. Knowledge on ICT ethics is important, however, sensitivity is more important. Sensitivity is defined formally in the next section.

*Corresponding Author: Woochun Jun, Seocho Dong, Seocho Gu, Seoul, Korea, wocjun@snue.ac.kr

The purpose of this work is to develop information and communication ethics sensitivity measurement standards for elementary school students. For this purpose, first, the initial sensitivity measurement standards are developed. Then, the proposed standards are surveyed by 30 elementary school teachers majoring information education. Finally, 19 sensitivity measurement standards with 6 areas are developed.

The rest of this paper is organized as follows. In Chapter 2, related works are introduced. In Chapter3, information and communication ethics sensitivity measurement standards for elementary school students are developed. In Chapter 4, statistical analysis is presented for those measurement standards. Finally, in Chapter 5, conclusions and further research issues are presented.

2. Related Works

2.1. Definitions of Sensitivity

Sensitivity is not formally defined in ICT ethics education. However, sensitivity is used for other ethics such as nursing ethics education. In nursing ethics education, moral sensitivity is defined as follows: Moral sensitivity is the ability to sensitize and detect conflict by sensitively identifying the problem surrounding the patient [7,8].

In this paper, sensitivity in information and communication ethics is defined as follows: Sensitivity is the ability to judge responsibilities for information communication ethics sensitively and recognize the consequences of their actions

Sensitivity is more important than knowledge on ICT ethics since knowledge does not relate with actions. That is, someone knows danger of copyright violation, however, he or she can keep doing misconducts.

2.2. Literature Review

In the literature, there are some works on awareness measurement standards of ICT ethics for students[9,10,11,12,13]. However, these works are concerned with awareness on knowledge. It means that the awareness measurement standards do not reflect sensitivity.

In this paper, some of awareness standards are introduced as follows.

In [11], information and communication awareness measure standards are proposed. The standards are developed based on TPB(Theory of Planned Behavior) that is widely used to explain the major variables affecting a person's behavior. The standards consist of 4 areas(moderation, respect, responsibility, and participation) and 57 detailed standards.

In [12], a measurement tool is developed for test consciousness of information and communication ethics for middle school students. The tool consists of 5 areas, respect, justice, etiquette, freedom, and responsibility, respectively. The tool also has final 33 standards after statistical analysis. The following Table 1 shows areas and description for each area in their work.

In [13], ICEI (Information and Communication Ethics Index) is proposed. They insist that, if the purpose of information and communication ethics education is to prevent and remedy the behavior of the information and communication ethics so that the youth themselves can judge the ethical and unethical in the information service society and do the right thing, then the development of the information and communication ethics index, which can measure the level of behavior of the information communication ethics problem of the youths, is more important than anything else. To ensure the reliability and feasibility of index developed in this study, it was applied to middle and high school students for an empirical analysis.

Table 1. Areas and Description of the Measurement Tool[12]

Area	Description
Respect	It is a tribute to the noble values of people and things in the information and communication society. This evaluates personal information protection and the perception of oneself and others as individuals.
Justice	This is a fair and correct way of doing things in the information and communication society. It evaluates the extent to which rules and laws are allowed to act properly and fairly.
Etiquette	This is the language etiquette and greeting etiquette that should be observed in the information and communication society. This evaluates the extent to which communication etiquette is observed and others are sought.
Freedom	This is a self-regulating and proactive attitude that does not harm others in the information and communication society. This evaluates the degree of self-control and self-regulation of one's will.
Responsibility	This is, in information and communication society, responsibility for its actions. It evaluates the extent of the act of disseminating information and the individual's own rules.

3. Development of Evaluation Metrics for Learners

3.1. Design Principles

The sensitivity measurement standards are developed with the following principles.

First, in order to develop comprehensive and integrated standards, overall aspects of education elements in ICT ethics education are considered and reflected.

Second, sensitivity is emphasized rather than knowledge and action for developing standards in this work. Sensitivity reflect interest and will to stop unfair action for elementary school students.

Third, too detailed standards such as standards depending on the current popular smart devices are not considered. This is

because too detailed standards will be changed or removed depending on popularity of the device or technology.

3.2. The Original Sensitivity Measurement Standards

The initial standards are developed with 6 areas and 19 standards. The following Table 2 shows the original standards.

Table 2. The Original Selection Standards

Area	Standards
General	1. I am interested in information and communication ethics. 2. I am willing to practice the recommended acts of ethics in everyday life. 3. I will stop immediately if I know of any harmful behavior in ICT.
Netiquette	4. I'm interested in netiquette 5. I am willing to practice the netiquette's recommended behavior in everyday life. 6. I will stop immediately if I know of any harmful behavior on netiquette.
Personal Information Protection	7. I am interested in privacy. 8. I am willing to practice the recommended acts for the personal information protection in my daily life. 9. I will stop immediately if I know of my personal information violation.
Copyright Protection	10. I am interested in copyright protection 11. I am willing to practice the recommended acts for the copyright protection in my daily life. 12. I will stop immediately if I know of my copyright violation
Internet Addiction	13. I am interested in internet addiction. 14. I am willing to practice the recommended actions to prevent internet addiction in my daily life. 15. I will stop immediately if I know of an act or habit that could lead to internet addiction. 16. I will be treated and consulted if it turns out to be an internet addiction.
Cyber Crime	17. I am interested in cybercrime. 18. I have a willingness to practice free behavior from cybercrime in my daily life. 19. I will stop immediately if I know my behavior is a cybercrime.

4. Statistical Analysis of the Proposed Standards

4.1. Overall procedure

In order to test the reliability of the proposed standards, extensive statistical works are done. First, 30 elementary school teachers in Seoul Metropolitan area in Korea are surveyed. Their majors are all information education. They are surveyed on May and June 2019. They are supposed to answer 5 scales for each standard as follows.

-Very likely(5 point)

-Slightly likely(4 point)

-Average(3 point)

-Slightly not likely(2 point)

-Very not likely(1 point)

4.2. Item Analysis

Table 3 shows the mean, standard deviation, skewness, and kurtosis of the 19 standards to measure students' sensitivity to information and communication ethics.

Table 3. Results of Item Analysis

No	M	SD	Skewness	Kurtosis
1	3.93	1.01	-0.50	-0.88
2	4.27	0.94	-1.11	0.28
3	4.37	0.96	-1.82	3.80
4	4.07	1.14	-1.18	0.59
5	4.37	0.93	-1.38	1.03
6	4.33	1.03	-1.56	2.21
7	4.27	0.91	-0.87	-0.42
8	4.17	0.91	-0.64	-0.79
9	4.27	1.05	-1.35	1.48
10	4.27	0.91	-0.87	-0.42
11	4.27	0.87	-0.91	-.0.08
12	4.30	0.99	-1.58	2.77
13	4.47	0.90	-1.72	2.18
14	4.47	0.82	-1.50	1.63
15	4.57	0.90	-2.67	8.11
16	4.20	0.92	-0.99	0.19
17	3.87	1.14	-0.63	-0.98
18	4.00	1.11	-0.80	-0.69
19	4.17	1.12	-1.31	1.02

Where M means mean and SD means standard deviation

As a result of the analysis, the mean ranged from 3.87 to 4.57 and the standard deviation ranged from 0.82 to 1.14. In addition, as a result of analyzing the skewness and kurtosis to examine the distribution of the items, the skewness range of each standard was -0.50□ -2.67 and the kurtosis range was -0.08□ 8.11. In general, if the absolute value of the skewness index is greater than 3.0, the skewness is severe, and if the absolute value of the kurtosis index is greater than 10, it can be determined that there is a problem in the normal distribution. Given these criteria, the descriptive statistics and item distribution of the items to measure students' evaluation of information and communication ethics sensitivity are good.

4.3. Reliability Analysis

In order to verify the reliability of the items measuring the sensitivity of information and communication ethics, the results

of analyzing the correspondence between the items using the Cronbach α value are shown in Table 4.

Table 4. Results of Reliability Analysis

Area	Item No.	Correlation Coefficient	Removal Factor
General	1	0.10	0.76
	2	0.57	-0.19
	3	0.37	0.33
Cronbach's α = 0.51			
Netiquette	4	0.44	0.84
	5	0.72	0.53
	6	0.62	0.62
Cronbach's α = 0.75			
Personal Information Protection	7	0.73	0.80
	8	0.78	0.75
	9	0.69	0.85
Cronbach's α = 0.86			
Copyright Protection	10	0.66	0.88
	11	0.88	0.69
	12	0.70	0.85
Cronbach's α = 0.86			
Internet Addiction	13	0.53	0.91
	14	0.89	0.78
	15	0.89	0.77
	16	0.66	0.87
Cronbach's α = 0.87			
Cyber Crime	17	0.62	0.91
	18	0.87	0.67
	19	0.73	0.81
Cronbach's α = 0.86			
(Overall) Cronbach's α = 0.94			

As shown in Table 4, Cronbach α showed 0.51 for general, netiquette 0.75, personal information protection 0.86, copyright protection 0.86, internet addiction 0.87, cyber crime 0.86, and overall information ethics sensitivity of 0.94, all above 0.50. The reliability is very good.

Table 5. Results of Correlation Analysis

Area	Correlation Value
General	0.796***(0.000)
Netiquette	0.859***(0.000)
Personal Information Protection	0.824***(0.000)
Copyright Protection	0.881***(0.000)
Internet Addiction	0.845***(0.000)
Cyber Crime	0.781***(0.000)

*** p<.001

4.4. Correlation Analysis

Table 5 shows the correlation between the information communication ethics sensitivity subdomain and the information communication ethics sensitivity.

Overall information and communication ethics (r = .796, p <.001), netiquette (r = .859, p <.001), personal information protection (r = .824, p <.001), copyright protection (r =. 881, p <.001), internet addiction (r = .845, p <.001), and cyber crimes (r = .781, p <.001) showed a statistically significant correlation with the overall information communication ethics sensitivity. Therefore, it can be seen that the overall information and communication ethics, which measure the sensitivity of general information and communication ethics, and netiquette, personal information protection, copyright protection, internet addiction and cybercrime are stable.

5. Conclusions and Further Research Issues

In the modern knowledge and information society, everyone is benefiting from advanced information techniques. However, various side effects also become a big problem. These side effects are time consuming and expensive to prevent and treat. Therefore, prevention through school education is very important.

At present, a variety of information and communication ethics education is provided to young adolescents. In addition, various measures have been developed to properly educate youth. Most scales, however, measure knowledge or will of them. Those scales do not reflect how they are sensible for information and communication ethics.

In this paper, the comprehensive measurement standards for sensitivity are developed. After the initial standards are developed, those standards are finalized through rigorous statistical analysis. Finally, 19 standards with 6 areas are proposed. These standards will be very helpful for measuring information and communication ethics sensitivity for elementary school students.

The immediate research issues are as follows. First, usability of these standards need to be tested by applying them to elementary school students. Second, it is necessary to develop extensive standards for middle school students and high school students, and adults. Recently, ethical issues for artificial intelligence are getting attention in the 4th industrial revolution era[14,15]. In the long run, the proposed sensitivity measurement standards need to be extended for artificial intelligence and big data issues.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgement

This work was supported by the 2019 Research Fund of Seoul National University of Education

References

[1] P. Song, H. Ryou, "A Study on VR Curriculum", Journal of Korea Internet of Things Society, **2**(4), pp. 21-25, 2016. https://doi.org/10.20465/KIOTS.2016.2.4.021

[2] G. Yoo, S. Kim, "Effects of STEAM activities using VR.AR on creative problem solving ability and peer interactions of children", The Journal of

Korea Open Association for Early Childhood Education, **24**(2), pp. 525-560, 2019, http://dx.doi.org/10.20437/KOAECE24-2-20

[3] Y. Chung, "Development of VR Fire-extinguishing Experience Education Contents Using UX Design Methodology", The Journal of the Korea Contents Association, **17**(3), pp. 222-230, 2017, http://dx.doi.org/10.5392/JKCA.2017.17.03.222

[4] J. Park, J. Han, "Development of Educational Contents in Robot Ethics for Elementary Students", Journal of Creative Information Culture, **4**(2), pp. 143-151, 2018.

[5] E. Yoo, M. Cho, "The Ethics of Robots and Humans in the Post-Human Age", The Journal of the Korea Contents Association", **18**(3), pp. 592-600, 2018, https://doi.org/10.5392/JKCA.2018.18.03.592

[6] S. Byun, H. Shin, J. Jeong, H. Kim, "A Study on Necessity of The Chart of Robot Ethics and its Contents", Journal of Ethics, **1**(112), pp. 295-319, 2017, http://dx.doi.org/10.15801/je.1.112.201703.295

[7] M. Yoo, K. Shon, "Effects of Nursing Ethics Education on Biomedical Ethics Awareness, Moral Sensitivity and Moral Judgment of Nursing Students", Journal of the Korean Bioethics, **12**(2), pp. 61-76, 2012.

[8] M. Park, D. Kjervik, J. Cradell, M. Oermann, "The Relationship of Ethics Education to Moral Sensitivity and Moral Reasoning Skills of Nursing Students", Nursing Ethics, **19**(4), pp. 568-580, 2012, https://doi.org/10.1177/0969733011433922

[9] J. Lee, K. Han, Y. Lee, S. Kim, "Development of an Index Model on the Information and Communication Ethics", Journal of the Korean Society of Computer Education, **10**(3), pp. 19-29, 2007.

[10] S. Cho, S. Kim, "Analysis of Factors to Measure Information Communication Ethics of Adolescents Using Delphi and AHP", Journal of the Korean Society of Computer Education, **11**(9), pp. 1-9, 2008.

[11] S. Cho, "The Development of Scale based on TPB to Measure the Information Communication Ethics Awareness", The Journal of Curriculum and Evaluation, **15**(1), pp. 291-320.

[12] I. Cho, J. Kim, "Measurement Tool Development for the Consciousness of Information and Communication Ethics", The Journal of Educational Information and Media, **15**(4), pp. 107-128, 2009.

[13] K. Kim, S. Kim, "Development and Validation of the Information and Communication Ethics Index of Adolescents", The Journal of Curriculum and Evaluation, **17**(3), pp. 191-221, 2014.

[14] V. Dignum, "Ethics in Artificial Intelligence: Introduction to the Special Issue", Ethics and Information Technology, **20**(1), pp. 1-3, 2018. https://doi.org/10.1007/s10676-018-9450-z

[15] A. Jobin, M. Lenca, and E. Vayena, "The Global Landscape of AI Ethics Guidelines", Nature Machine Intelligence, 1, pp. 389-399, 2019, https://doi.org/10.1038/s42256-019-0088-2

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 5, 174-185 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

DFIG Defects Diagnosis Method for Wind Energy Conversion Chain

Fatima El Hammouchi[1,*], Lamia El Menzhi[2], Abdallah Saad[1]

[1]National Higher School of Electricity and Mechanic, Hassan II University, Casablanca 8118, Morocco

[2]National School of Applied Sciences, Abdelmalek Essaadi University, Tangier 9000, Morocco

ARTICLE INFO

Article history:
Received: 18 August, 2019
Accepted: 23 September, 2019
Online: 08 October, 2019

Keywords:
Diagnosis
Defect
Doubly fed induction generator
Frequency spectrum
Fast Fourier Transform
Lissajous curve
Wind turbine

ABSTRACT

This paper is an extension of research work originally presented in 2018 IEEE fifth International Congress on Information Science and Technology (CiSt). The research consists on developing method to diagnose electrical defects affecting wind turbine doubly-fed induction generator DFIG which constitutes a crucial part of wind energy conversion chain. First off all, we create a model of a non-defected wind conversion system based on mathematical equations introduced in Matlab Simulink. Then, we apply an indirect vector control stator field orientation in order to increase wind energy performance. With the aim of diagnosing the defects attacking wind turbine generator, we propose a method based on grouping of fast Fourier transform spectral analysis and Lissajous curves performed to generator stator and rotor currents. This diagnosis technique is applied to wind turbine in normal operation (non-defected generator) in order to have a reliable reference data for asynchronous generator behaviour. However, connected to the grid, wind turbine generator is affected by various faults occurring in electrical power networks. Therefore, the diagnosis method is applied also to a defected generator. Considering diversity of grid defects, we deal in the current paper with open stator supplying phases and open rotor feeding phases due to rotor side converter legs opening. Indeed, this diagnosis method allows diagnosing generator defects type and severity by comparing the resulting frequency spectrum analysis and Lissajous curves under abnormal condition operating to reference data obtained in case of non-defected generator. So, our proposed method contributes to DFIG defects identification and anticipation. The simulations had been accomplished using Matlab Simulink. These results proved the efficiency and effectiveness of the proposed DFIG diagnosis method for wind energy conversion chain.

1. Introduction

To satisfy the world increasing energy demand we need to add new capacity to the grid while reducing global CO2 emissions and conserving our environment. Without doubt renewable energy in general and wind power in particular can form pillar supporting economic development and growing energy needs.

Indeed, the 51.3 Gigawatts of new installations bring total cumulative installations up to 591 Gigawatts by the end of 2018 [1]. This worldwide capacity showed that wind sector continues to progress, draining investments in wind energy.

Morocco take advantage of the latest technology in wind field and increase installed wind power to 2 Gigawatts by 2020 [2]. Actually, the overwhelming majority of Moroccan wind farms house horizontal axis wind turbines. For these wind turbines variable speed industry, the doubly fed induction generator DFIG is the most used [3]. Thanks to its excellent operational and control features [4], the DFIG makes wind energy integration with electrical networks easy and effective. In fact, much research focused on how to improve the integration of wind energy in the grid. This problematic have double face.

On one hand, wind power is intermittent energy source. So, integration of the wind energy can unfortunately affect the power system negatively. Furthermore, the electrical grids have been conventionally conceived for unidirectional energy flows from power stations going to cities. The use of dispersed wind energy

* Fatima EL HAMMOUCHI, National Higher School of Electricity and Mechanic, Hassan II University, Casablanca 8118, Morocco, Contact: +212 668 085 468, Email: elhammouchi.fatima@gmail.com

generators is more likely to result in bi-directional flow and may exacerbate problems with voltage and fault management [5]. Hence, in order to inject wind power to the grid, wind turbines generators must have the ability to contribute to both the voltage and frequency control in stabilising the power system following disturbance [6].
On the other hand, wind turbines generators must be connected to reliable distribution networks. So, grid operators need to ensure that electric networks operate continuously in a safe way [7]. But, faults in electrical power networks are inevitable and unpredictable events. As, wind turbines are connected to grid, they are consequently exposed to various grid faults. Unfortunately, faults in electrical grids can immediately paralyze electricity production in wind farms and sometimes lead to a loss of total structure. Actually, when grid fault occurs wind turbine is disconnected in order protect the structure and reconnected when normal operation has been resumed. However, for large wind energy capacity disconnection from the grid could generate problems in the control of frequency and voltage in the system [8]. Thus, wind turbines generators must be able to avoid excess fault levels while still contributing to fault identification and stop their propagation to save money and time. For this reason, it is necessary to know behaviours of DFIG wind turbine when grid faults are detected.

Consequently, the effectiveness DFIG defects diagnosis remains a necessity.

In the current paper a diagnosis method is developed for wind turbine doubly-fed induction generator. This method is based on analysis of frequency spectrums and Lissajous curves of DFIG stator and rotor currents.

Firstly, we present the wind energy chain modeling based on mathematical equations created in Matlab Simulink.

Then, in order to increase performance of wind system, we chose to control stator flux by the application of an indirect field oriented control (IFOC) to DFIG. Afterwards, we use stator and rotor currents to establish frequency spectrums analysis and to draw Lissajous curves. This diagnosis method was applied in the first phase to a non-defected generator in order to have a reliable reference data in healthy case of DFIG. The second phase consists on using the method while generator is under most known abnormal condition operating due essentially to grid faults.

In the last part, the resulting Lissajous curves and frequency spectrums in case of defected generator are compared to Lissajous curves and frequency spectrums obtained when generator is under normal operating conditions (reference data: presented in previous article).

This method allows diagnosing wind turbine generator defects type and severity; thereafter, anticipating serious damage affecting wind energy conversion chain.

2. Wind Energy Chain Conversion Modeling

2.1. Wind Turbine Modeling

The essential components of wind turbine installed in Moroccan parks are presented in the figure bellow. The modeling is about horizontal-axis wind turbines HAWT using DFIG. In fact, the wind turbine generator is asynchronous machine doubly-fed: the stator and rotor are both connected to power supply. The DFIG stator is connected to national grid through protection contactor. The rotor is connected to the grid using the transformer and converters "back-to-back" composed of rotor side converter (RSC) and grid side converter (GSC). A capacitor is placed between two converters to reduce the voltage ripples. Indeed, regardless of the rotor power magnitude and direction, the GSC has to keep the dc-link capacitor voltage at a set value and to guarantee a converter operation with unity power factor (zero reactive power) [9].

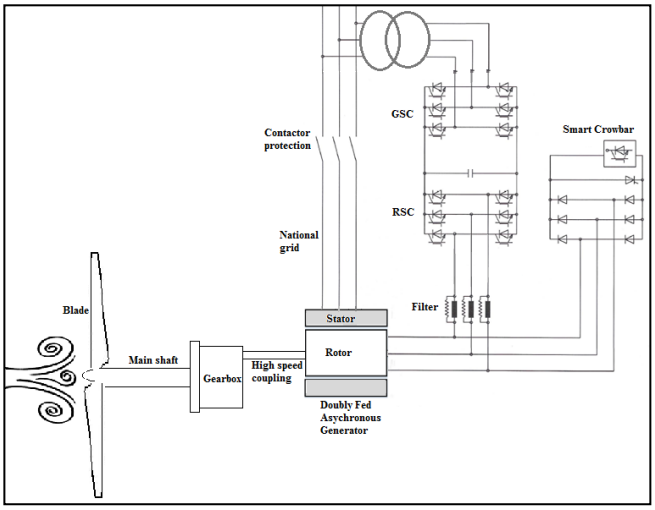


Figure 1: Main Components of Wind Turbine.

We introduced sum of different sinusoidal signals: varied amplitudes and frequencies in order to create wind speed model in Matlab Simulink as showed in the figure 2, which reflects approximately a recorded wind speed in Moroccan wind farms [10].

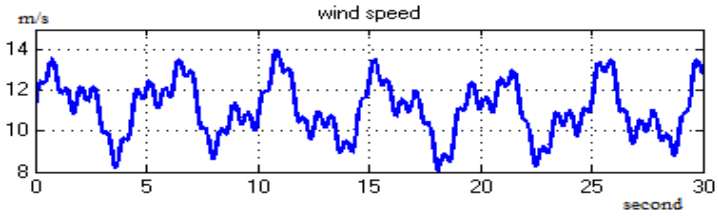


Figure 2: Wind Speed Model (m/s).

We assume that frictions are neglected and with the use of equations presented in previous paper [11], the wind turbine Matlab Simulink model is created.

Also, the DFIG modeling is based on equations of the stator and the rotor voltages (Us, Ur) in Park reference (d, q) as presented below:

$$\begin{cases} U_{ds} = \dfrac{d\varphi_{ds}}{dt} + R_s.i_{ds} - \varphi_{qs}\omega_s \\ U_{qs} = \dfrac{d\varphi_{qs}}{dt} + R_s.i_{qs} + \varphi_{ds}\omega_s \\ U_{dr} = \dfrac{d\varphi_{dr}}{dt} + R_r.i_{dr} - \varphi_{qr}(\omega_s - \omega_r) \\ U_{qr} = \dfrac{d\varphi_{qr}}{dt} + R_r.i_{qr} + \varphi_{dr}(\omega_s - \omega_r) \end{cases} \quad (1)$$

With:

$$\omega_r = \omega_s - P.w \quad (2)$$

In park reference, the flux equations can be expressed as follows:

$$\begin{cases} \varphi_{ds} = L_s.i_{ds} + M.i_{dr} \\ \varphi_{qs} = L_s.i_{qs} + M.i_{qr} \\ \varphi_{dr} = L_r.i_{dr} + M.i_{ds} \\ \varphi_{qr} = L_r.i_{qr} + M.i_{qs} \end{cases} \quad (3)$$

$$L_s = l_s - M_s, \qquad L_r = l_r - M_r$$

The electromagnetic torque is expressed according to equation below:

$$T_{em} = p(\varphi_{ds}.i_{qs} - \varphi_{qs}.i_{ds}) \quad (4)$$

Where **p** represents pole pairs number of the DFIG.

We chose to control DFIG by stator flux. We seek to cancel the indirect stator flux in order to keep that direct stator flux. In this mode of DFIG operating conditions, control of the torque and flux is decoupled. This method is applied to improve the DFIG active and reactive power control performance, which allows controlling flux and torque dynamically and separately.

Based on equations (3) and (4) we have:

$$\begin{cases} i_{ds} = \dfrac{1}{Ls}(\varphi_s - M.i_{dr}) \\ i_{qs} = -\dfrac{M}{Ls}.i_{qr} \\ T_{em} = p\varphi_s.i_{qs} = -p\dfrac{M}{Ls}\varphi_s.i_{qr} \end{cases} \quad (5)$$

In general, for large wind turbine, the generator stator resistance is negligible.

Also, it is supposed that the electrical network is stable and the field is constant. Thus, by using equation (1) we get:

$$\begin{cases} U_{ds} = 0 \\ U_{qs} = \varphi_s\omega_s \end{cases} \quad (6)$$

So, the DFIG active and reactive powers of can be written as equations below:

$$\begin{cases} P_s = U_{ds}.i_{ds} + U_{qs}.i_{qs} \\ Q_s = U_{qs}.i_{ds} - U_{ds}.i_{qs} \end{cases} \quad (7)$$

We get the expression of the active and reactive powers:

$$\begin{cases} P_s = -U_s\dfrac{M}{Ls}.i_{qr} \\ Q_s = \dfrac{U^2{}_s}{\omega_s.L_s} - U_s\dfrac{M}{Ls}.i_{dr} \end{cases} \quad (8)$$

So, the equation (8) shows that the control of the active power is independent of the reactive power. Indeed, the active power is controlled by iqr the quadrature rotor current whereas the reactive power can be controlled by idr the direct rotor current.

$$\begin{cases} U_{dr} = (R_r + S(L_r - \dfrac{M^2}{Ls})).i_{dr} - g\omega_s(L_r - \dfrac{M^2}{Ls}).i_{qr} \\ U_{qr} = (R_r + S(L_r - \dfrac{M^2}{Ls})).i_{qr} + g\omega_s(L_r - \dfrac{M^2}{Ls}).i_{qr} + g\dfrac{M}{Ls}.U_s \end{cases} \quad (9)$$

Also, to perform the machine control power there are two approaches for the field oriented control direct and indirect:

- Direct field oriented control is simple to implement : it consist on ignoring the terms of coupling in equation (9) and to establish an independent regulator in each axis to control the active and reactive power independently. But the direct field oriented control is not the most efficient [12];
- Indirect field oriented control: in this control we consider the coupling terms in equation (9) and we look for compensating them to control the powers and rotor currents. The indirect field oriented control is being more frequently used as it guaranties the easy operation over all the speeds range [13].

Table 1: Symbols and Abbreviations for generator Modeling

Symbol	Quantity
φ_s **(d, q)**	Stator flux components in Park reference
φ_r **(d, q)**	Rotor flux components in Park reference
Rs	Stator resistance
Rr	Rotor resistance
Ls	Stator cyclic inductances
Lr	Rotor cyclic inductances
ls	Stator leakage inductances
lr	Rotor Leakage inductances
Ms, Mr	Mutual inductances between stator and rotor phases respectively
M	Maximum mutual inductance between stator and rotor stage (the axes of the two phases coincide)
i_s ***(d,q)***	DFIG stator current components in Park reference
i_r ***(d,q)***	DFIG rotor current components in Park reference
Tem	Electromagnetic torque
U_s ***(d,q)***	Stator voltage components in Park reference
Ur ***(d,q)***	Rotor voltage components in Park reference
P_s	DFIG active power
Q_s	DFIG reactive power

For DFIG modeling, we chose in this study to orient the stator flux of generator according to the direct axe. Moreover, we use the indirect vector field oriented control (IFOC) to increase the system performance. The wind energy conversion chain model created in Matlab Simulink is illustrated in figure 3 below.

The figure 3 presents the model of wind turbine created in Matlab Simulink which is composed of:

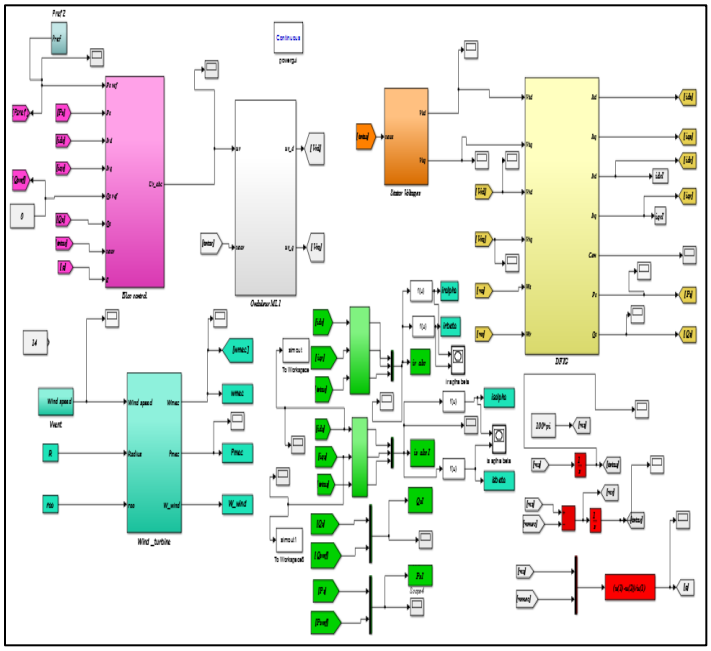


Figure 3: Wind Turbine Model in Matlab Simulink.

- The blue bloc at the bottom in the left contains modeling equations of wind turbine. In order to have maximum C_p we take $\beta = 0$. Also, the wind speed model introduced as illustrated in figure 2;
- The yellow bloc at the top on the left includes the DFIG stator and rotor flux and currents equations. This bloc is linked to the power supply (orange bloc);
- The chosen control of generator is implemented in the pink bloc at the top on the right. Actually, the DFIG indirect field oriented control allows us to control dynamically and separately flux and torque. The bloc is connected to the grey bloc of Pulse Width Modulation PWM;
- We use the Concordia transformation in order to obtain DFIG currents Lissajous curves as figured in green blocs;
- We use mathematic equations in red blocs in order to get the DFIG slip g.

2.2. Simulations and Results

The graphs below represent the simulation at the first thirty seconds of the wind turbine model created in Matlab Simulink. We suppose that frictions are neglected.

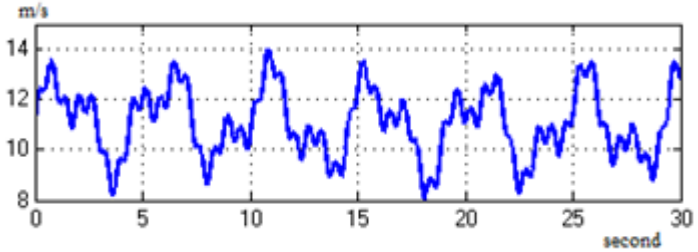


Figure 4: Wind Speed (m/s).

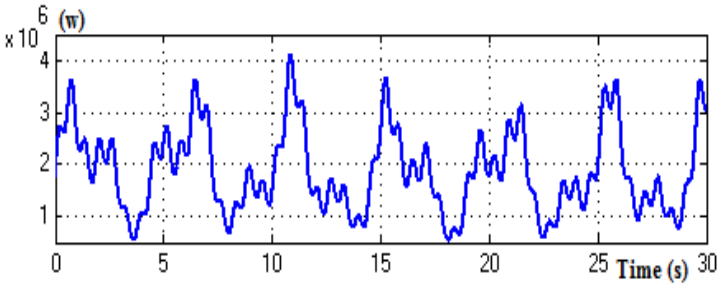


Figure 5: Power Captured by Blade (w).

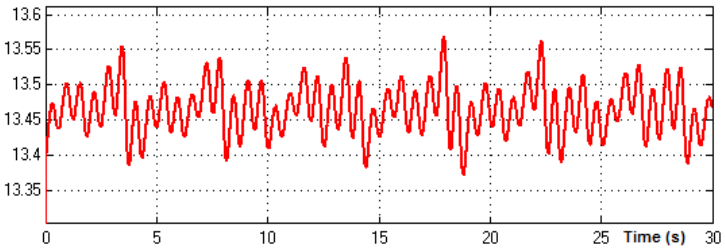


Figure 6: Specific Speed λ.

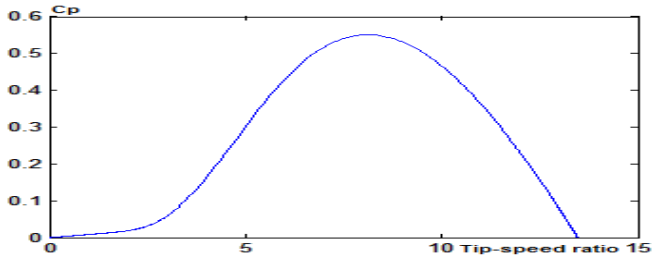


Figure 7: Cp Coefficient of Performance at β =0.

In fact, figure 4 represents the variable wind speed which we apply to wind system conversion. The figures 5 and 6 show the system performance. Actually, the power captured by blade follows the same pace of wind speed.

The figure 7 shows coefficient of performance as a function of specific speed λ and the angle of orientation of the blade β. The coefficient reaches its maximum value which is 0.56 when β is zero and λ is optimal. Using the indirect vector field-oriented control (IFOC), we got the following figures representing asynchronous generator performance.

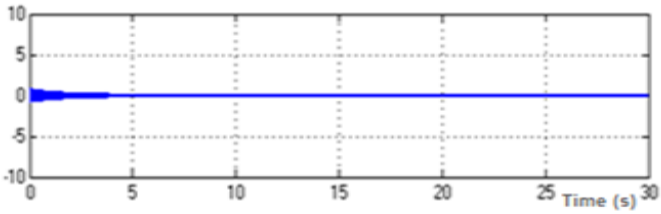


Figure 8: DFIG Indirect Stator Field φ_{qs} (wb).

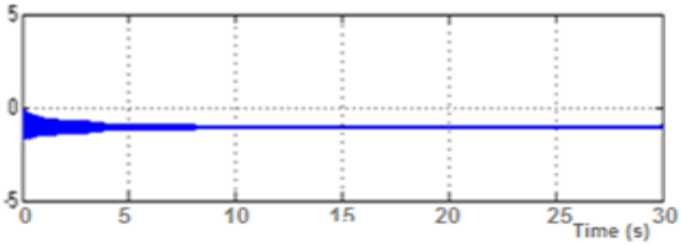


Figure 9: DFIG Direct Stator Field φ_{ds} wb.

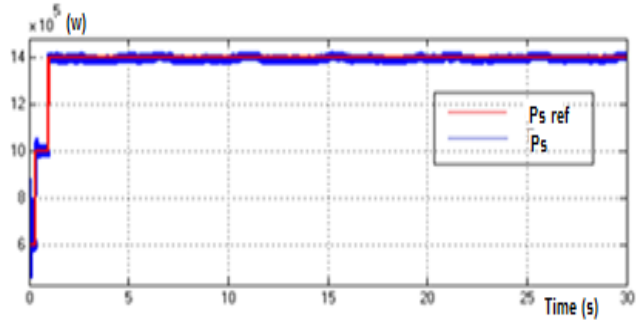


Figure 10: Comparison between Ps and Psref.

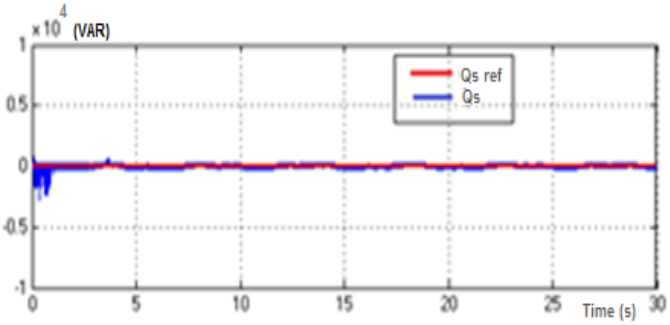


Figure 11: Comparison between Qs and Qsref.

In figures 8 and 9 the indirect field values of stator machine reached to zero and the direct stator field is constant since the stator flux of DFIG is chosen to be oriented following the direct axe. Furthermore, we choose Psref a reference for active power as a three steps form applied in different time. As illustrated in figure 10, the real active power pursues Psref. In addition, as showed in figure 11, the system reactive power follows Qsref which around zero value.

A zoom is presented in figure 12 of doubly-fed asynchronous generator three-phase stator currents.

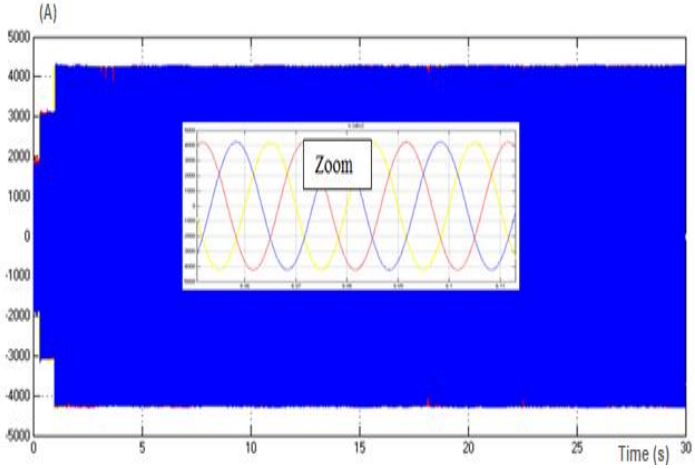
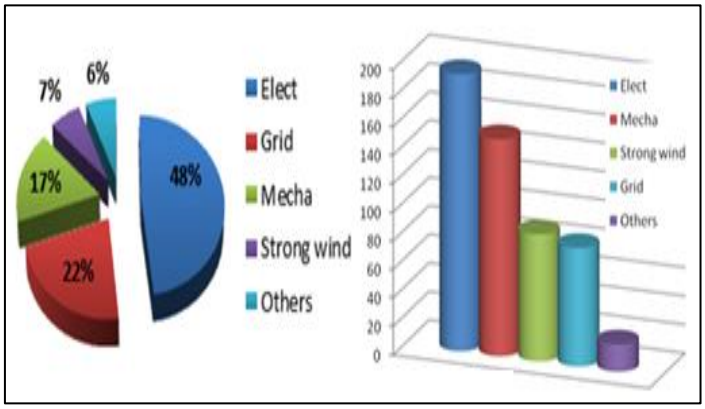


Figure 12: DFIG Three Phase Stator Currents.

It is shown that DFIG three-phase stator currents isabc have exactly sinusoidal shapes.

3. DFIG Diagnosis Defects Method

Nowadays, the electricity generated from wind turbines keeps increasing. However, faults in electrical power networks are inevitable events. According to data collected from yearly defects study of wind turbines Moroccan park [14], wind turbine connected to grid is exposed to various grid faults: 22% of defects are due to grid faults as presented in figure 13 (a). Also, grid defects can cause significant losses to wind system even if grid defects frequency remains less than other defects frequencies figure 13 (b) .



(a) (b)

Figure 13: Wind Turbines Defects: (a) type and (b) frequency.

The figure 13 shows defects type and frequency according to yearly Moroccan park data defects inspection.

Indeed, as indicated in sector graph figure 13.a wind turbines suffer from various internal electrical faults (48%), we mentioned for example over-current in DFIG stator phases, overvoltage in crowbar, overcurrent phase in GSC, thermal tripping in yaw drive motor because of cable isolation failure, etc . Also, linked to the electrical network, wind turbines are exposed to several grid faults (22%) such as lack of grid voltage, variation in the voltage magnitude and time duration and overheated generator due to the short circuit in the power supply, etc. The third defects type which represents 17% of total defects types is mechanical defects that occurs mainly in gearbox, electrical generator bearings, yaw motor, and also blades bearings. The rest defects types are essentially due to strong wind and other types.

The histogram above in figure 13.b shows different defects frequencies. It proves that the electrical defects are the most recurrent: they occur more than 194 times during a year in this park. In the second place come the mechanical defects with 152 times, followed by strong wind (89 times), grid faults (83 times) and others less than 20 times.

Usually, when fault occurs in the grid wind turbine is disconnected to avoid loss structure; however, for large wind energy capacity disconnection from the grid could generate problems. For this reason, multitudes of techniques are used to monitor the wind energy conversion chain health continually [15, 16].

In this current paper we propose diagnosis method for wind turbine DFIG electrical defects. The method is based on association of Fast Fourier Transform introduced and Lissajous curves analysis.

Indeed, Fast Fourier Transform (FFT) allows a conversion of a signal from its original domain to a representation in the frequency domain. The FFT analysis is applied in this paper to generator

stator currents in order to determine the harmonic components amplitude and to get the maximum amplitude spectrum of the wavelet coefficients. The resulting spectrum offers a source of defects information [17].

To FFT analysis we associate Lissajous curves analysis. Currently, in the field of Lissajous curves, recent developments are realized in different domains such as the rotating machinery field [18], and defects diagnosis [19]. The method to get Lissajous curves of stator and rotor currents was explained in detail in previous paper [20]. Thanks to the curves obtained we will have an idea about defects kind impacting generator in wind conversion energy chain. In fact, when the defects appear, the Lissajous curves will be deformed depending on defects kind.

In the next part diagnosis method will be applied firstly to a non-defected generator.

4. Diagnosis Method Applied in Case of DFIG Normal Condition Operating

In order to have a reliable reference data for behaviour of DFIG in healthy case, the diagnosis method is developed for generator during defects-free operation.
As showed in figure 14, the FFT diagram of the stator three-phase current is implemented in Matlab Simulink.

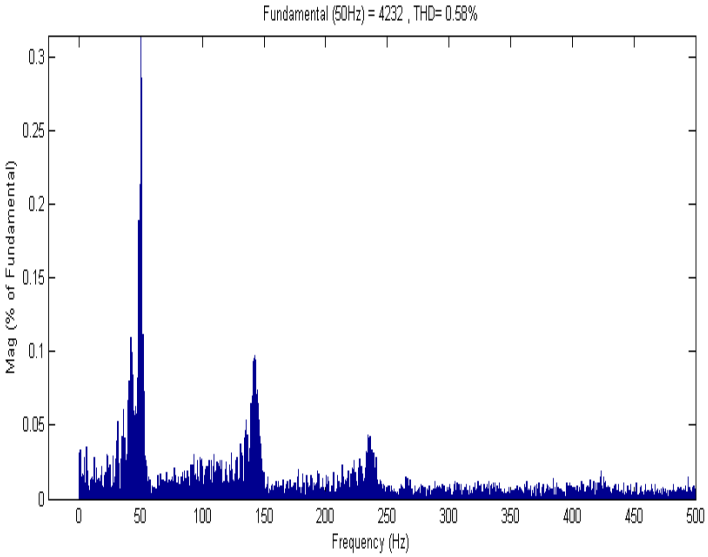


Figure 14: Fast Fourier Transform Spectrum of Stator Currents.

The signal fundamental frequency is at 50 Hz and the Total Harmonic Distortion (THD) which measures harmonics content is equal to 0.58%. In addition, harmonics are detected at frequency equals to 150 Hz and at 250 Hz. These harmonics are due to the use of PWM Pulse Width Modulation in the system, they represent the third and fifth harmonics. Also, magnitude percentages are respectively 0.1% and 0.05%.

Also, Lissajous curves of DFIG stator and rotor three-phase currents are simulated in Matlab Simulink.

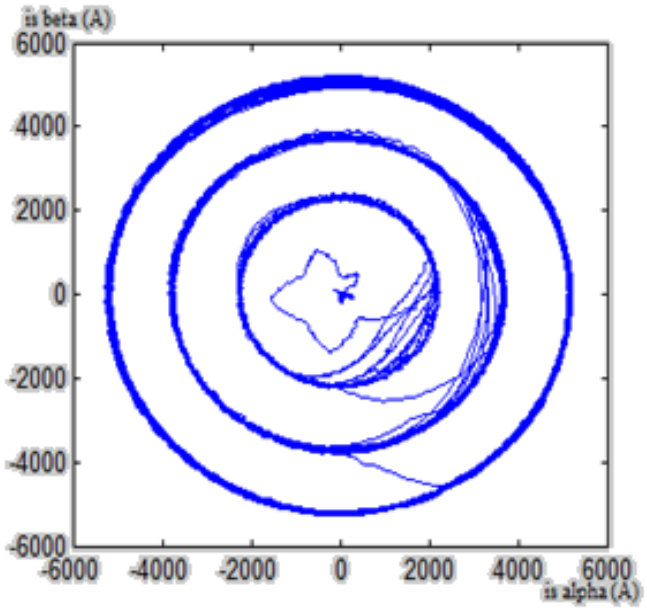


Figure 15: DFIG Stator Currents Lissajous Curves.

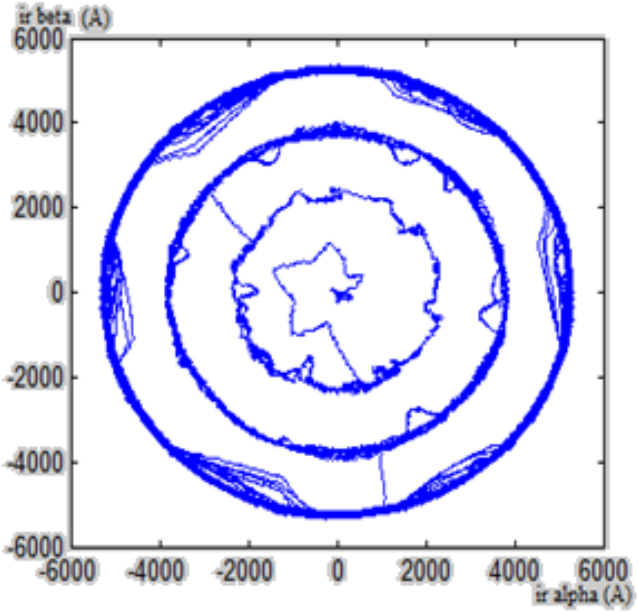


Figure 16: DFIG Rotor Currents Lissajous Curves.

The figures 15 and 16 illustrate DFIG stator and rotor currents Lissajous curves when generator is under normal operating condition. In this case, Lissajous curves have the form of three circles. These circles are the result of reference active power Psref which had been chosen as a signal of three steps presented in figure 10. Also, the rotor currents Lissajous curves have six petals flower shape (figure 16). This shape is explained by the fact that the rotor generator power supply is realized through the inverter with 6 switches.
In the next part we apply diagnosis method to defected DFIG.

5. Diagnosis Method Applied in Case of Open Stator Feeding Phases

As mentioned previously, larger wind turbines were developed and installed in windy world parts. Currently, the wind power

generated is comparable to conventional fossil-fuel energy.
So, the impact of grid faults on wind power plants is considerable and could lead to blackouts if we allowed disconnection of wind turbines from grid. To keep wind turbines connected to the grid despite the presence of faults, we need to monitor and diagnose continuously the structure. Indeed, wind turbine generator is subjects to grid faults [21, 22]. We mentioned lack or variation of the voltage grid magnitude and frequency, and short circuit. Some recurrent grid faults are diagnosed by using the proposed diagnosis method [23].
Considering diversity of grid defects, we deal in this part with the problem of open stator feeding phases. This defect occurs rarely. However, when it occurs, it mainly affects generator and can cause significant losses.
In fact, when the stator is fed from two phases, the generator will continue to operate but the phase imbalance will cause its heating which can lead to eventual damage to DFIG, thereafter the loss of the whole structure.

5.1. Open stator supplying phase 'a'

The figure 17 represents FFT of stator currents in case of generator fed from two phase power supply (phases 'b' and 'c').

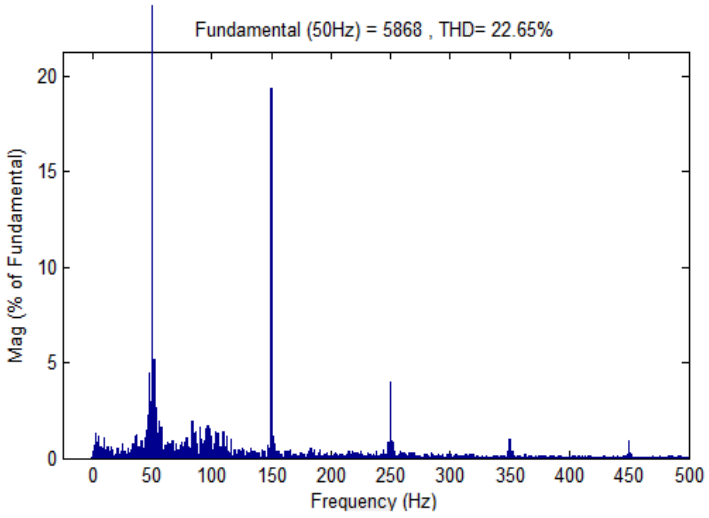


Figure 17: Stator Currents FFT Spectrum in Case of Open Stator Supplying Phase 'a'.

In addition to harmonics at frequency equals to 150 Hz and at 250 Hz (case of defects-free generator), peaks appear in new frequency (at 25 Hz and 75 Hz). Also, Total Harmonic Distortion THD passed from 0.58% to 22.65% in case of defected generator. Indeed, higher THD means increase of generator heating and peak currents. As showed in figure 9 fundamental magnitude was 0.32%. But, when one feeding stator phase 'a' was opened, this fundamental magnitude value increases and passed to 24%.

The following figures show Lissajous curves respectively generator stator and rotor currents in case of open stator supplying phase 'a'.

Comparing stator and rotor Lissajous curves obtained in cases of defects- free and defected generator, we see clearly in figures 19 and 20 that stator and rotor curves are deformed and the currents values increase.

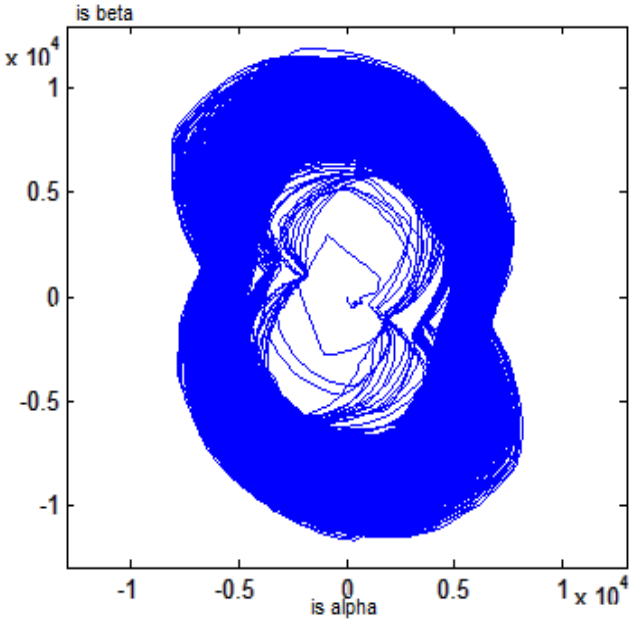


Figure 18: DFIG Stator Currents Lissajous Curves in Case of Open Stator Supplying Phase 'a'.

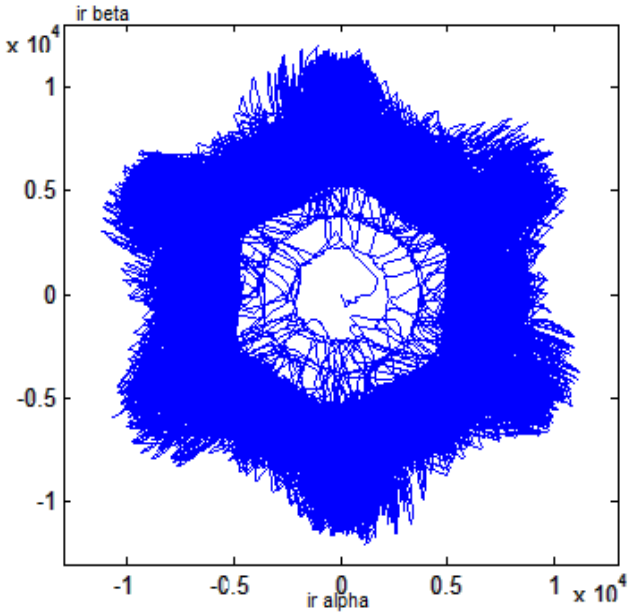


Figure 19: DFIG Rotor Currents Lissajous Curves in Case of Open Stator Supplying Phase 'a'.

In fact, the opening of one stator feeding phase (phase 'a') generates an increase of the currents in the other phases. Also, the stator currents beta reach higher values than alpha currents. The curves have now elliptical shapes according to the vertical axis (stator currents beta). Also, curves circumference and thickness increased, which makes difficult observation of three separate curves.

The same for rotor currents Lissajous curves showed in figure 19 they are superimposed on each other because of increasing curves thickness. Which makes impossible detection of the three separate curves as Lissajous curves obtained in case of normal condition

operating. Also, curves have six petals flower shape but they are deformed comparing to the reference curves.

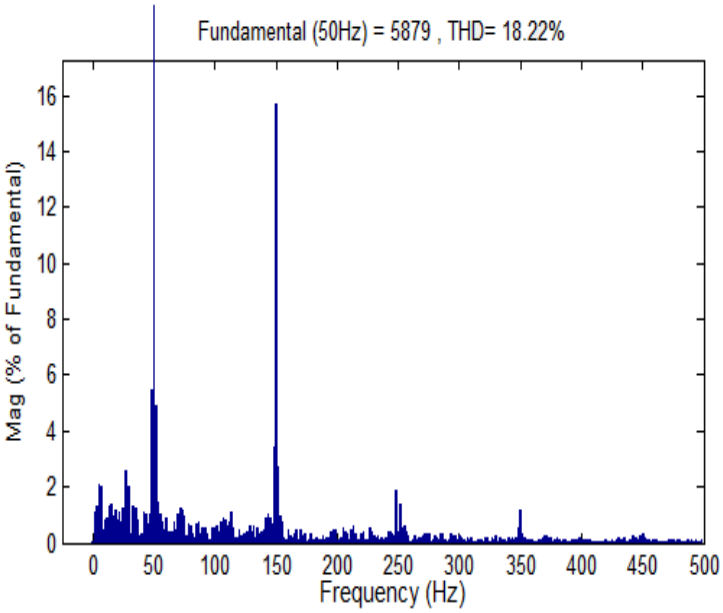


Figure 20: Stator Currents FFT Spectrum in Case of Open Stator Supplying Phase 'b'.

5.2. Open stator supplying phase 'b'

As illustrated in figure 20, the Fast Fourier Transform spectrum of DFIG stator currents when supplying phase 'b' is opened has approximately the same magnitude of fundamental in case of opened phase 'a'. Also, harmonics at frequency 150 Hz, 250 Hz and 350 Hz appear in the FFT spectrum.

In term of Total Harmonic Distortion THD in this case, the value is 18.22%, a little less comparing to open supplying phase 'a'.

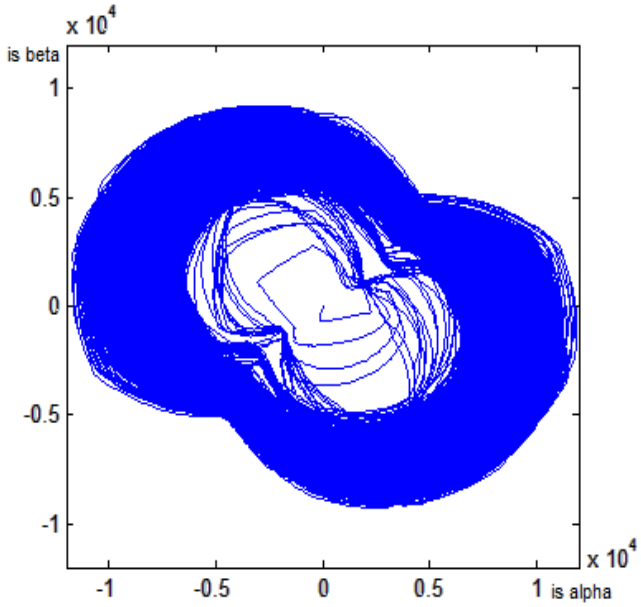


Figure 21: DFIG Stator Currents Lissajous Curves in Case of Open Stator Supplying Phase 'b'.

In case of open stator feeding phase 'b', stator and rotor Lissajous curves obtained respectively in figures 21 and 22, are deformed comparing to the case of opening phase 'a'.

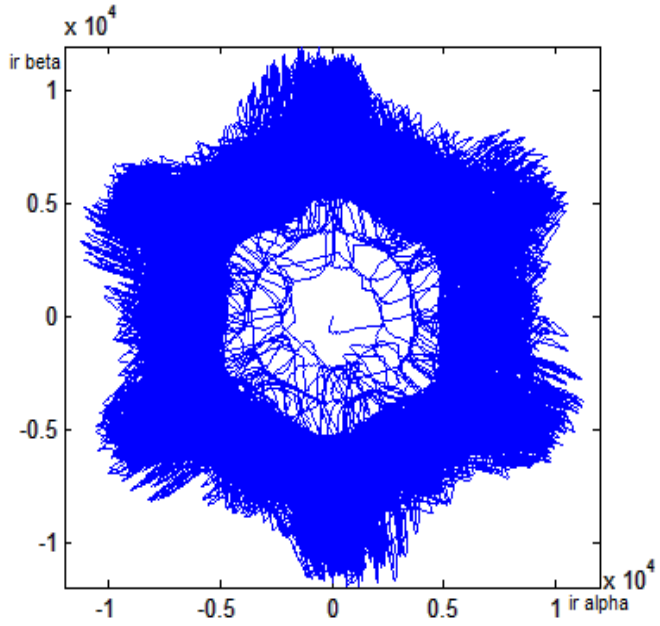


Figure 22: DFIG Rotor Currents Lissajous Curves in Case of Open Stator Supplying Phase 'b'.

Indeed, the stator currents curves have elliptical shapes. But, for open phase 'b' stator currents alpha reach higher values than stator currents beta. Also, we note an inclination angle of $2\pi/3$ of Lissajous curves compared to open phase 'a'.

For rotor currents Lissajous curves showed in figure 22, we note that the curves have the same shape as figure 19 case of open phase 'a' with a slight increase in rotor currents.

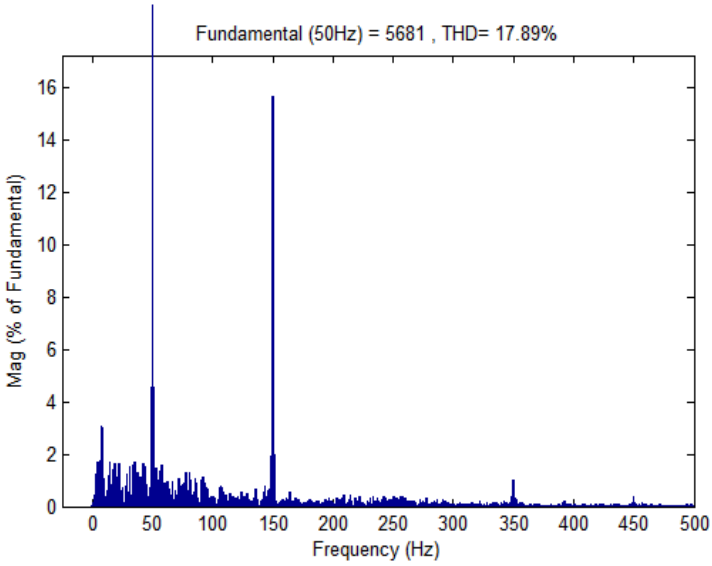


Figure 23: FFT Spectrum of Stator Currents Case of Open Stator Supplying Phase 'c'.

5.3. Open stator supplying phase 'c'

As showed in figure 23, the FFT spectrum of stator currents in the case of open supplying phase 'c' has approximately the same magnitude of fundamental in case of open phase 'a' or phase 'b'. In addition, the THD in the case of open supplying phase 'c' is 17.89 %, it has approximately the same value comparing to open supplying phase 'b'.

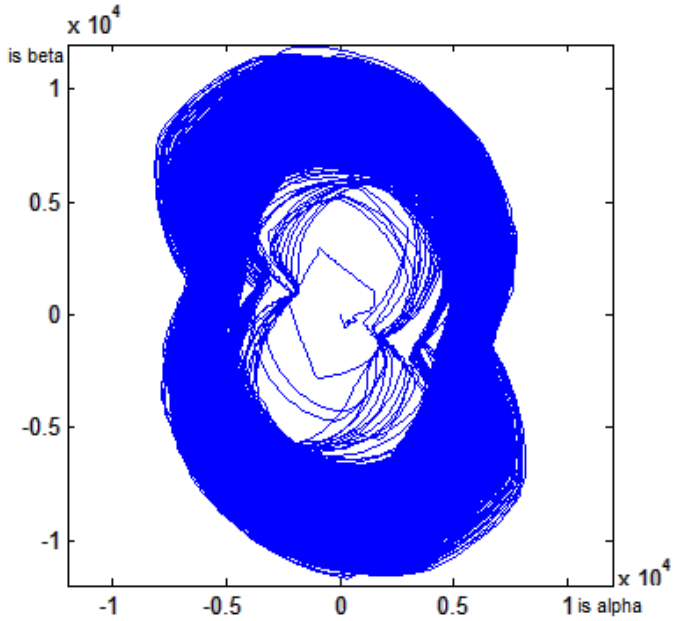


Figure 24: DFIG Stator Currents Lissajous Curves in Case of Open Stator Supplying Phase c.

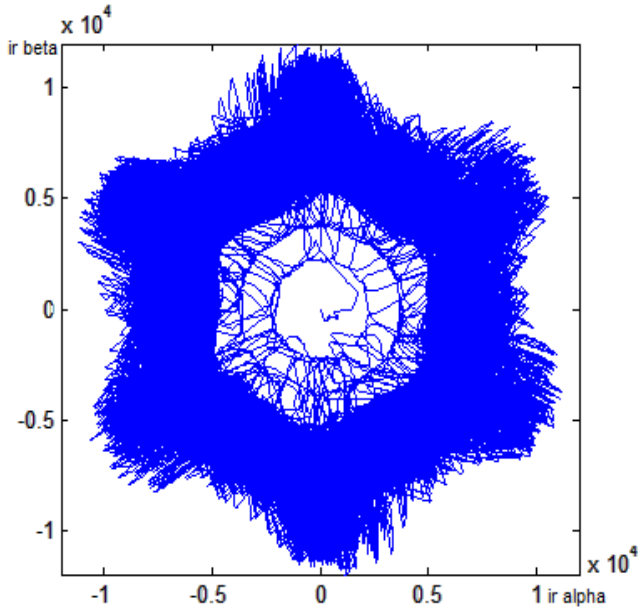


Figure 25: DFIG Rotor Currents Lissajous Curves in Case of Open Stator Supplying Phase 'c'.

The figures 24 and 25 represent in that order stator and rotor Lissajous curves. Curves are deformed comparing to reference Lissajous curves the case of defects-free operation. Also, the stator beta currents reach higher values than alpha currents as curves when phase 'a' is opened, but in this case the stator beta currents have higher values.

As showed in figure 25, rotor currents Lissajous curves have kept the same shape as case of open phase 'a' and phase 'b'.

We summarize that, the opening of one feeding DFIG stator phase generates an increase of the currents in the other phases and overheating of generator. Indeed, the resulting FFT spectrums analysing of stator currents help determine the amplitude of the harmonic components. These peaks amplitude, give us an idea about the defects severity affecting the wind energy chain.

Furthermore, the stator and rotor Lissajous curves are deformed according to defected generator supplying phase. So, we can detect DFIG defect types and know exactly which feeding phase is faulty.

6. Diagnosis Method Applied in Case of Defected Rotor Side Converter

Situated between the wind turbine generator and distribution networks, the back-to-back power converter has to satisfy both the generator side and grid side requirements [24]. As mentioned in paragraph 2.1, the control performed is indirect stator field oriented control. This way, DFIG is controlled by the rotor side converter. Owing to high voltage and a large transient current in the rotor windings, the power converter switching devices may be damaged due to their low power ratings of the switching devices [25]. So, connected to rotor circuit, a defected converter can cause generator damages. Hence, to diagnose DFIG rotor defects we apply a DFIG diagnosis method in case of open rotor side converter legs.

The figure below presents rotor side converter RSC components.

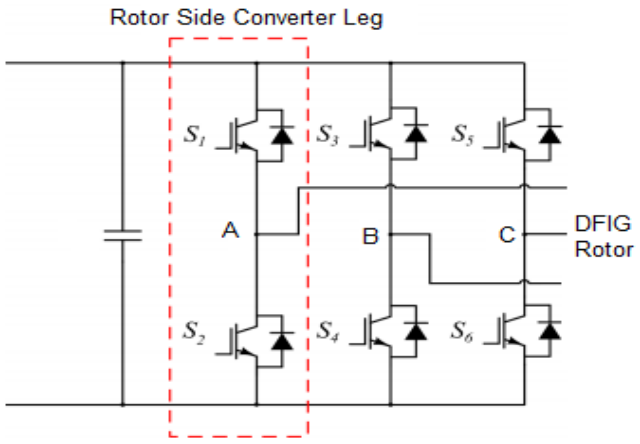


Figure 26: Rotor Side Converter with 3 Legs.

In the following part, we applied diagnosis method to DFIG in case of one rotor side converter leg opening.

The figures hereafter present the simulations results obtained in Matlab Simulink when RSC legs respectively (S1, S2), (S3, S4) and (S5, S6) are opened.

6.1. Open rotor feeding phase 'a'

The figure above exposes the FFT spectrum of stator currents in the case of opened RSC leg (S1, S2). The fundamental magnitude value reaches 4782, and THD is 89.45%. Comparing these values to reference frequency spectrum in the case of DFIG defects free operation, we note that the fundamental magnitude exceeded the reference fundamental magnitude. Also, we observe an absence of harmonics at 150 Hz and 250 Hz due to defected RSC.

In term of total harmonic distortion, his value passed from 0.58% to 89.45% in case of one open feeding rotor phase. This high THD level of harmonics in stator currents waveforms cause increasing of DFIG temperature which reduce generator life and subsequently lead to the loss of generator.

The following figures show Lissajous curves respectively generator

stator and rotor currents in case of open rotor feeding phase 'a'.

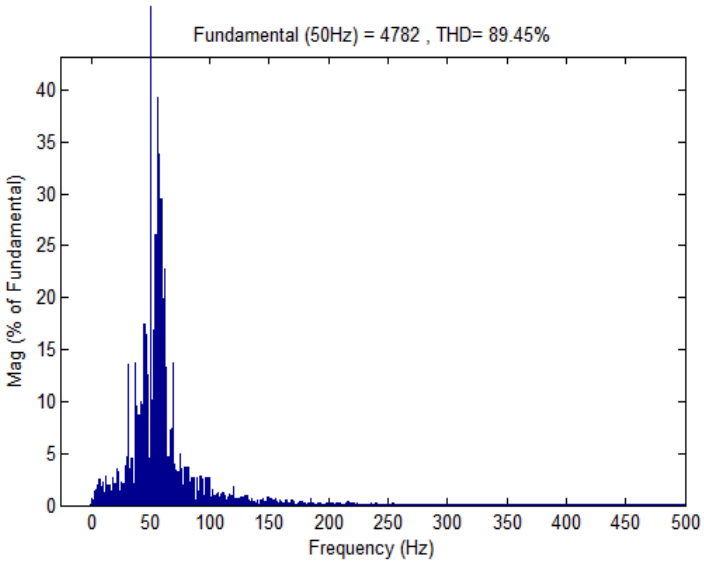


Figure 27: Stator Currents FFT Spectrum in Case of Open RSC Leg (S1, S2).

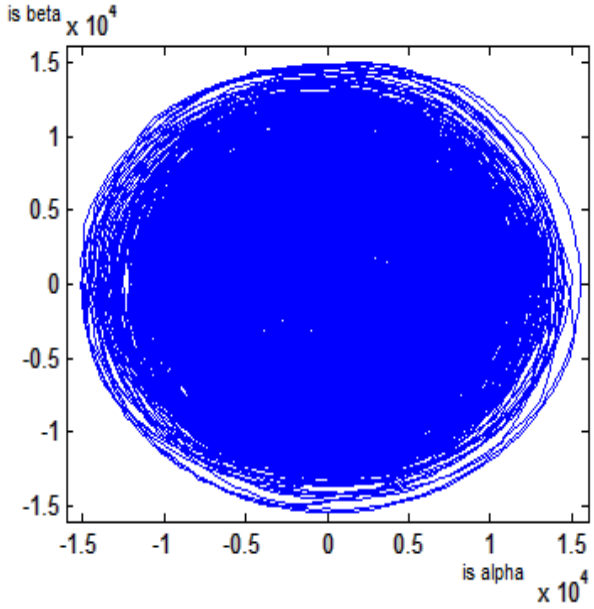


Figure 28: DFIG Stator Currents Lissajous Curves in Case of Open RSC Leg (S1, S2).

Comparing stator Lissajous curves obtained in cases of defects- free and defected generator, we see clearly in figure 28 that stator Lissajous currents curves have kept the circular shape whereas circumference and thickness increased because of stator currents increasing which makes harder the differentiation of three separate curves.

Lissajous rotor curves, as presented in figure 28, are totally deformed and the rotor currents increase. In fact, the opening of one rotor feeding phase (phase 'a') generated by opening of one rotor side converter (S1, S2) increase the currents in the other phases and cause the curves superimposition on each other. Which make impossible detection of the three separate curves as Lissajous curves obtained in case of generator under normal condition operating.

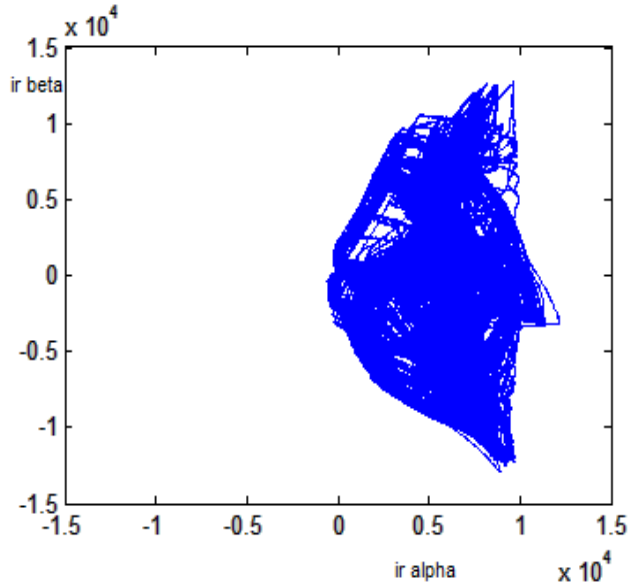


Figure 29: DFIG Rotor Currents Lissajous Curves in Case of Open RSC Leg (S1, S2).

Also, the curves have now two deformed petals flower shape instead of six petals obtained in reference data. The rotor currents ir alpha in this case are located on the positive part of Lissajous curves.

In the next part, we apply diagnosis method to DFIG in case of open RSC leg (S3, S4).

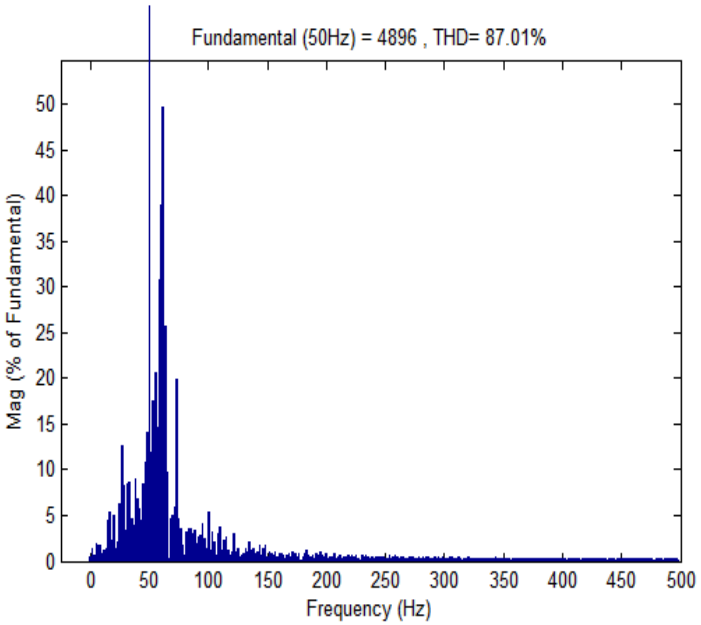


Figure 30: Stator Currents FFT Spectrum in Case of Open RSC Leg (S3, S4).

6.2. Open rotor feeding phase 'b'

The figure 30 presents the FFT spectrum of stator currents in the case of open rotor feeding phase 'b'. The fundamental magnitude reaches 4896 which is more than case of opened leg (S1, S2). The THD remains high with 87%. As the first case in paragraph 6.1 the harmonics at 150 Hz and 250 Hz did not appear in this frequency spectrum.

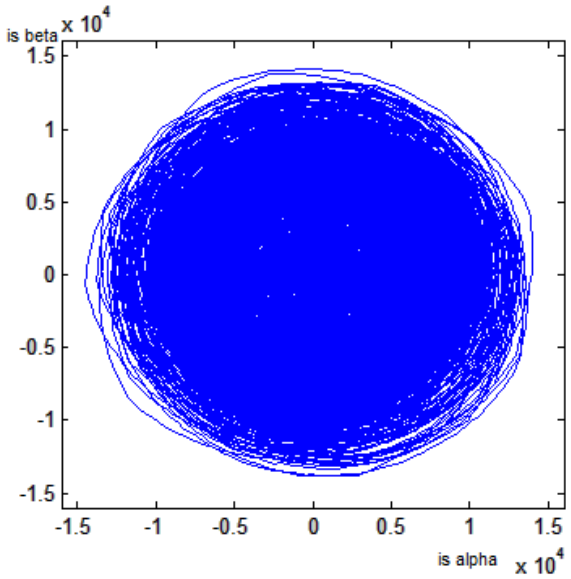


Figure 31: DFIG Stator Currents Lissajous Curves in Case of Open RSC Leg (S3, S4).

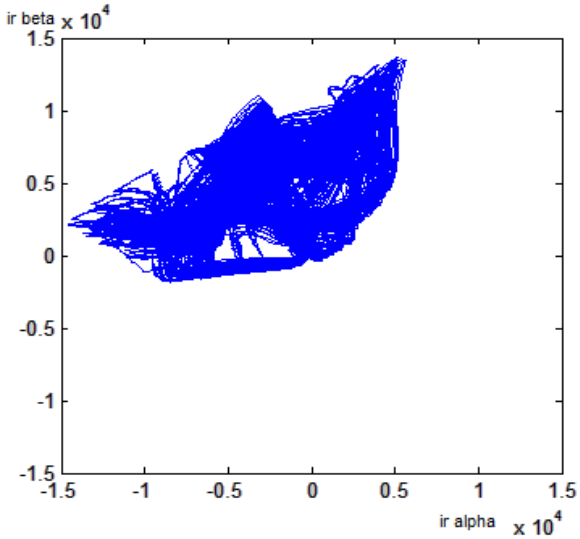


Figure 32: DFIG Rotor Currents Lissajous Curves in Case of Open RSC Leg (S3, S4).

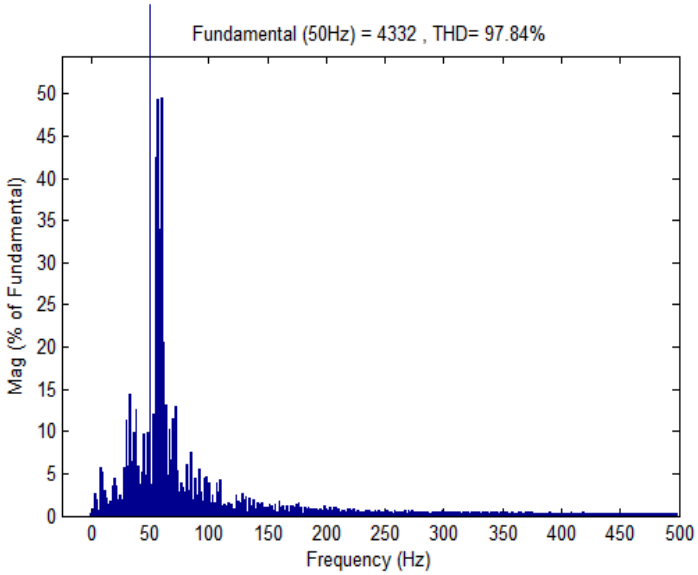


Figure 33: Stator Currents FFT Spectrum in Case of Open RSC Leg (S5, S6).

When rotor feeding phase 'b' is opened the stator Lissajous curves obtained in figure 31 have kept the circular shape. They are similar to the first defected rotor case.

For rotor currents, the opening RSC leg (S3, S4) as presented in figure 32, generates Lissajous curves with shape of two petals flower. But these two petals are inclined by $2\pi/3$ compared to the case of opening rotor phase 'a'. For the last part of this paper, our proposed DFIG diagnosis method is applied in case of open RSC leg (S5, S6).

6.3. Open rotor feeding phase 'c'

In the case of open rotor feeding phase 'c', we obtain the FFT spectrum of stator currents showed in figure 33. The fundamental magnitude reaches 4332 which is less than case of opened legs (S1, S2) and (S3, S4).
The THD reaches a maximum percentage with approximately 98% comparing to the previous opened phases.
According to the FFT spectrum graph and as foregoing cases, we note that the harmonics at 150 Hz and 250 Hz did not figure out.

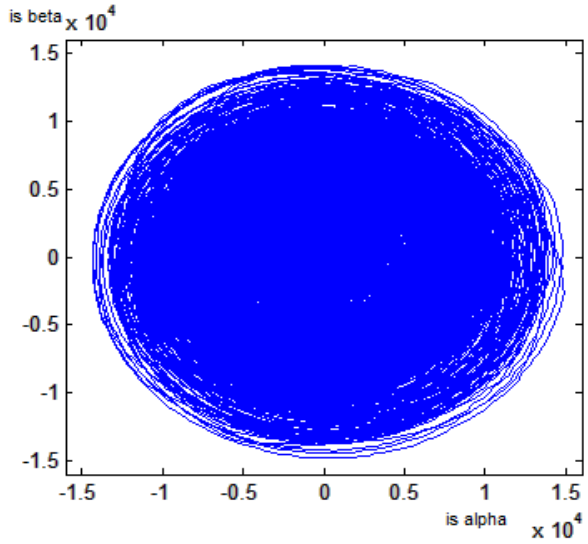


Figure 34: DFIG Stator Currents Lissajous Curves in Case of Open RSC Leg (S5, S6).

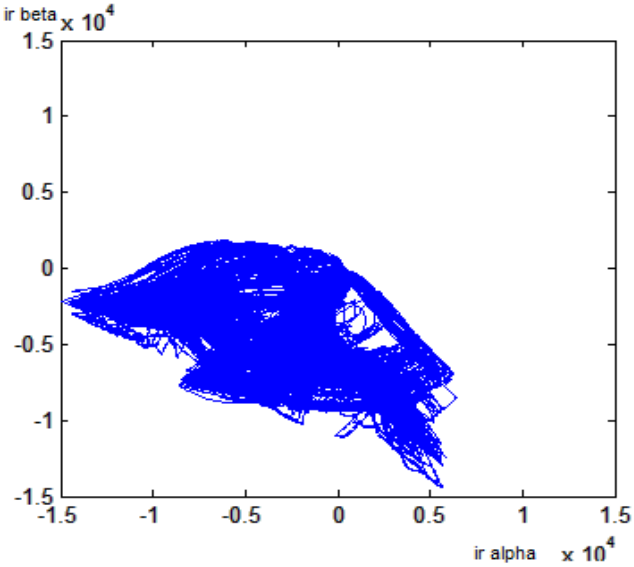


Figure 35: DFIG Rotor Currents Lissajous Curves in Case of Open IGBT Leg (S5, S6).

As opened feeding rotor phases 'a' and 'b', the stator Lissajous curves showed in figure 34 have similar shape. The generating Lissajous curves in figure 35 form two petals flower which are inclined by 4π/3 compared to the antecedent cases.

We deduct that, when one feeding DFIG rotor phase is opened the generator currents in the other phases increased. These high currents cause overheating of generator and can lead to the loss of generator. Through the frequency spectrums analysis of stator currents we determine the peaks amplitude, which show us the severity of the DFIG defects. And thanks to the stator and rotor Lissajous curves we detect exactly which feeding phase is faulty by means of deformation curves.

7. Conclusion

In this research paper, the method to diagnose wind turbine DFIG defects is developed through studying and analysing frequency spectrum and Lissajous curves of generator stator and rotor currents.

We start with creating wind energy conversion chain model in Matlab Simulink. Then, we apply indirect field oriented control to have separate and dynamical control on system flux and torque. Besides, the IFOC granted to have optimal, easy and efficient generator system operation over all the speed levels. Afterwards, based on studying generator currents frequency spectrum analysis and Lissajous curves, the method was developed with the aim of diagnosing and anticipating the potential DFIG wind turbine defects.

In the first step, this method was applied to generator under normal operating condition to get reliable reference data of generator behaviour.

In the second step, the technique was performed in case of open stator and rotor feeding phases. The results were compared with reference data obtained in case of defect-free operation.

Actually, according to defects types and by using Fast Fourier transform analysis, peaks appear or disappear at defined frequencies. Also, the maximum amplitude spectrum of the wavelet coefficients increasing, gives us an idea about the defects severity. Moreover, stator and rotor currents Lissajous curves are deformed considering shapes: thickness, circumference and direction of Lissajous curves in case of normal condition operating generator. We deduct that Lissajous curves are deformed according to the kind of defects attacking generator in wind turbine. Hence, the proposed DFIG diagnosis method allows recognition and prevention of potential defects attacking wind turbine.

Using Matlab Simulink, all the simulations had been realized. These results confirmed the efficiency and effectiveness of the proposed DFIG diagnosis defects method in the wind energy field.

References

[1] Global Wind Energy Council GWEC, "Global Wind Energy Report" April 2019.

[2] Moroccan Investment Development Agency, Kingdom of Morocco.

[3] R. M. R. Muthu, "Doubly Fed Induction Generator for Wind Energy Conversion System - a Survey", International Conference on Energy Efficient Technologies for Sustainability, pp. 617-628, 2013.

[4] M. A. El-Sharkawi, "Wind Energy an Introduction", University of Washington, Seattle, USA: 219, 2016.

[5] Outhred, Hugh & R. Bull, Stanley, "Meeting the Challenges of Integrating Renewable Energy into Competitive Electricity Industries", 2007.

[6] T.R. Ayodele, A.A. Jimoh, J.L Munda and J.T Agee, "Challenges of Grid Integration of Wind Power on Power System Grid Integrity: A Review", International Journal of Renewable Energy Research", Vol.2, No.4, 2012.

[7] K. Loudiyi, et al., "Grid Code Status for Wind Farms Interconnection in Northern Africa and Spain: Descriptions and Recommendations for Northern Africa", Renewable and Sustainable Energy Reviews, Elsevier, vol. 81, part 2, pp. 2584-2598, 2018.

[8] Hansen, et al., "Grid Fault and Design-Basis for Wind Turbines - Final Report", Roskilde, Risø National laboratoriet for Bæredygtig Energi, Denmark, Forsknings center Risoe. Risoe-R, no. 1714, 2010.

[9] M. Abomahdi and M.Srivastava, "Control of Grid Voltage and Power of Doubly Fed Induction Generator wind turbines during grid faults", IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE), vol. 9, iss. 4, pp 12-21, 2014.

[10] F. El Hammouchi, L. El Menzhi and A. Saad, "Wind Turbine Doubly-Fed Asynchronous Machine Diagnosis Defects-State of the Art", DEStech Transactions on Environment Energy and Earth Sciences, pp. 300-306, October 2017.

[11] F. El Hammouchi, L. El Menzhi, A. Saad, Y. Ihedrane and B. Bossoufi, "Wind Turbine Doubly-Fed Asynchronous Machine Diagnosis Defects -Part two", 2018 IEEE 5th International Congress on Information Science and Technology (CiSt), Marrakech, pp. 449-454, 2018. DOI: 10.1109/CIST.2018.8596498

[12] M. Allam, B. Dehiba, M. Abid , Y. Djeriri and R. Adjoudj, "Etude Comparative entre la Commande Vectorielle Directe et Indirecte de la Machine Asynchrone à Double Alimentation (MADA) Dédiée à une Application Eolienne", Journal of Advanced Research in Science and Technology (JARST),vol. 1, no. 2, pp. 88-100, 2014.

[13] Y. Ihedrane, C. El Bekkali, and B. Bossoufi, "Power Control of DFIG-Generators for Wind Turbines Variable-Speed", International Journal of Power Electronics and Drive System, vol. 8, no. 1, pp. 444-453, 2017.

[14] F. El Hammouchi, L. El Menzhi and A. Saad, "Wind Turbine Double-Fed Asynchronous Machine Diagnosis Defects-Part One", International Congress of Industrial Engineering and Systems Management CIGIMS, Meknes, Morocco, pp. 380-383, Mai 2017.

[15] A. Joshuva and V. Sugumaran, "Fault Diagnostic Methods For Wind Turbine: A Review", ARPN Journal of Engineering and Applied Sciences, vol. 11, no. 7, pp. 4654-4668, 2016.

[16] L. M. Popa, B. B. Jensen, E. Ritchie and I. Boldea, "Condition monitoring of wind generators", 38th IAS Annual Meeting on Conference Record of the Industry Applications Conference, 2003., Salt Lake City, UT, USA, vol. 3, pp. 1839-1846, 2003. Doi: 10.1109/IAS.2003.1257819

[17] E. ElBouchikhia, V. Choqueuse and M. Benbouzid,"Induction Machine Faults Detection Using Stator Current Parametric Spectral Estimation", Mechanical Systems and Signal Processing, vol. 52–53, pp. 447-464 , 2015.

[18] H. A. H. Al-Khazali and M. R. Askari, "Geometrical and Graphical Representations Analysis of Lissajous Figures in Rotor Dynamic System", Journal of Engineering (IOSR), vol. 2, issue. 5, pp. 971-978, 2012.

[19] L. El Menzhi, and A. Saad, "Lissajous Curve of an Auxiliary Winding Voltage Park Components for Doubly-Fed Induction Machine Electrical Faults Diagnosis", The Journal of Advanced Materials Research, vol. 860-863, pp. 2223-2231, 2014.

[20] F. El Hammouchi, L. El Menzhi, A. Saad, Y. Ihedrane, and B. Bossoufi, "Wind Turbine Doubly-Fed Asynchronous Machine Diagnosis Defects Using Stator and Rotor Currents Lissajous Curves", IJPEDS International Journal of Power Electronics and Drive System, Vol. 10, No. 2, pp. 961-970, 2019.

[21] A. Gupta , Dr. A. Shandilya, "Challenges of Integration of Wind Power on Power System Grid : a Review", International Journal of Emerging Technology and Advanced Engineering ,vol. 4, iss. 4, April 2014.

[22] A. Mouahid, A. Khamlich, A. Saad, "Study Of The Behaviour Of Wind Turbines Using Doubly Feed Asynchronous Generator During Voltage Dip", Mediterranean Telecommunication Journal, vol. 5, no. 2, pp. 126-132, june 2015.

[23] F. El Hammouchi, L. El Menzhi and A. Saad, "Diagnosis Method for Wind Turbine Doubly Fed Induction Generator under Grid Defects", International Journal of Information Science & Technology – IJIST, vol. 3, no. 3, pp. 46-55, 2019.

[24] F. Blaabjerg, M. Liserre, and K. Ma, " Power Electronics Converters for Wind Turbine Systems", IEEE Transactions on Industry Applications, vol. 48, no. 2, pp. 708-719, 2012

[25] J. J. Justo and F. A. Mwasilu, "Low Voltage Ride through Enhancement for Wind Turbines Equipped with DFIG under Symmetrical Grid Faults", Tanzania Journal of Engineering and Technology, vol. 37, no. 2, pp. 125-136, December 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 180-187 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

An Algorithm to Improve Data Accuracy of PMs Concentration Measured with IoT Devices

Mihaela Balanescu[1,*], George Suciu[1], Marius-Alexandru Dobrea[1], Cristina Balaceanu[1], Radu-Ioan Ciobanu[2], Ciprian Dobre[2], Andrei-Cristian Birdici[1], Andreea Badicu[1], Iulia Oprea[2], Adrian Pasat[1]

[1]*Beia Consult International, Research & Development, 041385, Romania*

[2]*University Politehnica of Bucharest, Faculty of Automatic Control and Computers, 060042, Romania*

ARTICLE INFO

Article history:
Received: 15 January, 2020
Accepted: 25 February, 2020
Online: 09 March, 2020

Keywords:
Air Quality
IoT
Edge/cloud computing
PMs concentration

ABSTRACT

Air pollution is responsible for increased morbidity and mortality due to respiratory problems mainly caused by long term exposure. Although the emissions of principal air pollutants are highly regulated, there is a lack of information about the real extent of personal exposure for an accurate health impact assessment. To tackle these challenges, local air pollution measurements and citizen involvement based on the small IoT devices became necessary. The Tel-MonAer platform is based on IoT devices and Edge/Cloud computing technologies and allows the (near) real-time monitoring of Particulate Matter air pollutants considering the complex chemistry and influence of various parameters (i.e. air humidity, wind speed, temperature). The aim of this paper is the assessment of the influence that air humidity has on the PM concentrations measured with IoT devices based on laser beam technologies. The results showed that in order to increase the accuracy of PM concentrations values a threshold value for relative humidity of 80% needs to be considered. When humidity values are below 80%, the PM concentration values are considered valid, while for values over the threshold, a specific correction algorithm needs to be applied. This paper presents the correction algorithm (based on the type of sensor and humidity) and the testing results (an increase of at least 2.5 times of the correlation coefficient between the corrected and reference values).

1. Introduction

This paper is an extension of the work originally presented in CSCS22: The 22nd International Conference on Control Systems and Computer Science, Bucharest, 2019 [1].

Worldwide, air pollution has extensive effects on the environment, human health and global economy, as research showed correlation between premature deaths and low air quality [2,3]. The extent of the consequences of air pollution levels are strongly related to the pollutant concentrations and the level of exposure. Until recently, the assessment of air quality has been strongly reliant on traditional monitoring networks, because of their accuracy, but they also have some disadvantages that should not be disregarded [4,5]. The main issues of these monitoring networks are high costs of acquisition, maintenance requirements, improper placement in areas with low pollution and the limited number of fixed stations, due to legal restrictions for location [6]. Therefore, the need for alternative air pollution measurements is indisputable, in the context of spatial variability of air quality [7-9]. As a result of the variety of sensors on the market, the increased computing power and new communication protocols and the community-led sensing initiative, the topic of air pollution became a key research topic, at local and regional scale [10].

The research community expressed concerns particularly regarding the dangerous effects on human health of two key pollutants: nitrogen dioxide and particulate matter (PM). The latter is one of the most dangerous pollutants in terms of health effects, as it can cause a wide range of negative reactions, even at low concentrations [11]. Among them, the PM_{10} (PM with diameter lower than 10 μm) and $PM_{2.5}$ (PM with diameter lower than 2.5 μm) are considered to have the greatest impact, as their effects are not only related to pollutant concentrations, but also to the frequency and the duration of exposure [12]. For individuals, there are also other factors that play important roles in the extent of air pollution effects, such as health status and age [13].

*Mihaela Balanescu, mihaela.balanescu@beia.ro

The prime sources of particulate matter in the atmosphere are either natural, such as volcano eruptions and forest fires, or human-made, such as traffic, industry, agriculture, construction and other combustion processes. PM concentrations are particularly important to monitor due to the fact that they can be emitted not only from direct emission sources, but also from chemical reactions between different gases, such as NOx and SO2 [14]. A comprehensive characterization of PM has to consider multiple factors: (1) mass; (2) elemental composition; (3) water-soluble ionic species; and (4) organic compounds. The traditional sampling systems based on gravimetric measurements of collected particles generate direct measurements of airborne particle mass. Moreover, during the sampling process, there is the possibility of losing the semi-volatile organic compounds and semi-volatile ammonium compounds (such as NH_4NO_3). The composition of the sample of PM is also decisive for the accuracy of the measurements, because the presence of ionic species (i.e. sulfate and nitrate compounds) increases the liquid water uptake of suspended particles and therefore, the particle dimension. Therefore, the chemical composition of the sample and the temperature heavily influence the correct assessment of PM concentrations in the atmosphere [15].

This paper presents an analysis of the variation of particulate matter (PM_{10} and $PM_{2.5}$) concentrations in relation to relative humidity. Chapter 2 compiles related work for data accuracy of PMs, Chapter 3 discusses the method that it is used, Chapter 4 presents the results, and lastly, Chapter 5 concludes the paper.

2. Related Work

The effects of different parameters on the data accuracy of PM concentrations were approached in several papers. The influence of wind and precipitation on different-sized particulate matter concentrations were investigated in paper [16], showing that the effects of atmospheric conditions differ, depending on the size of the particulate matter. The increase in wind speed can decrease the concentrations of fine PM, while decreasing the concentrations of coarse PMs. The authors also found a stronger negative impact of precipitation on PM_{10} than on $PM_{2.5}$.

In paper [17], authors analyze the way PM_{10} concentrations are influenced by different meteorological parameters, such as pressure, relative humidity, temperature, wind speed, wind direction, CO, SO_2, NO, NO_2. A quantile regression model has been employed and the results showed that the influence of the independent variables was significant in at least one or more quantiles of the PM_{10} concentrations. Among the analyzed parameters, relative humidity was proven to have a significant impact on quantiles 0.05 to 0.3 and an insignificant impact at higher quantiles.

The topic of the relationship between relative humidity and PM concentrations was approached in paper [18]. Authors found that PM concentrations in the atmosphere are closely correlated with the levels of relative humidity. It has been shown that high humidity conditions (between 70-100%) led to a reduction in $PM_{2.5}$ concentrations, while low-humidity conditions (below 70%), led to the increase in $PM_{2.5}$ concentrations. In case of PM_{10} concentrations, humidity values below 45% had an accumulation effect, causing an increase in concentration, while an environment with humidity levels above 45% led to lower concentrations.

3. Methods

3.1. Tel-MonAer platform

The aim of the Tel-MonAer project was the development of a mobile, extensible and scalable system which integrates technologies such as the Internet of Things and Edge/Cloud Computing, for the purpose of monitoring and performing real time analysis of the risk factors of public health and the environment. The architecture of the IoT platform is presented in Figure 1.

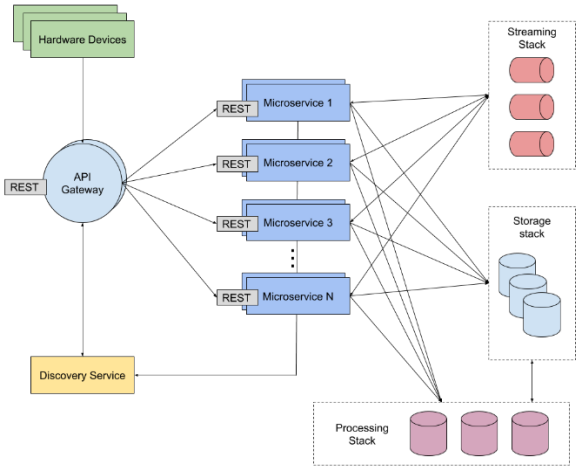


Figure 1: NETIoT architecture



(a).BEIA Consult, Romania.



(b).University Politehnica of Bucharest, Romania

Figure 2: Installed air-quality sensors that send the data the Tel-MonAer platform.

The platform registers every hardware device employed by the end-user, such as sensors or device gateways, with a unique ID. The data from the sensors is firstly received by the MQTT protocol, through a device gateway and then forwarded to a cloud gateway. The Tel-MonAer system is capable of simultaneously monitoring different air quality parameters such as SO_2, NOx, CO, O_3, PM_{10}, $PM_{2.5}$, as well as meteorological parameters (wind direction and speed, pressure, temperature, relative humidity).

The air quality data is currently being collected from IoT sensors in two locations: the premises of BEIA Consult International and University Politehnica Bucharest, as shown in Figure 2.

Tel-MonAer is designed to allow some specific features like availability and scalability. Moreover, the platform will permit further development. The architecture of the platform is based on microservices, because of the advantages of this model, such as independent, faster and more cost-effective development of each microservice and dedicated and specific databases for each component.

The high volumes of data stored by the Tel-MonAer platform demand a scalable and performant storage layer. For this purpose, Apache Cassandra database has been used because of its ability to scale almost linearly, to tackle failover situations and to automatically replicate data in more data centers.

The data is further processed by the platform, using two types of processing. Batch processing is used for analyzing the data received from multiple sensors and within a specific time frame and for performing predictions of possible evolutions. Real-time processing is used for event detection. We used Apache Spark, a general use engine for both real-time and batch processing, because of its advantages, such as in-memory processing, real-time stream processing and sophisticated analytics support.

3.2. Details of the method used

The parameters of the data set used to perform the analysis are: PM10 and PM2.5 concentration values, atmospheric pressure, atmospheric temperature and relative humidity. The measurements were performed in Bucharest using Libelium sensors. The parameters were measured between the 1st of November 2018 and the 28th of January 2019, with a frequency of 15 minutes.

The process of data acquisition follows several steps: accessing the gateway interface, connecting to the MySQL database interface to access the sensor data, logging into the phpMyAdmin interface, querying the database for hourly average values, downloading the data selected by the query function.

4. Experimental Results

The dataset resulted from the registered measurements contains 2133 values for every parameter. Firstly, a qualitative analysis of the data has been performed, in order to compare the measured values with standard data requirements. Secondly, a preliminary analysis has been carried out using statistical descriptive methods for the parameter, such as variation, mean value and standard deviation [1].

The variation of PM_{10} and $PM_{2.5}$ concentrations function of relative humidity is presented in Figure 3 and Figure 4, respectively. The results show an increase in PM concentrations for values of relative humidity greater than 90%. This is a strong indication of a measurement error, caused by the measurement method or by the complex chemistry of PMs.

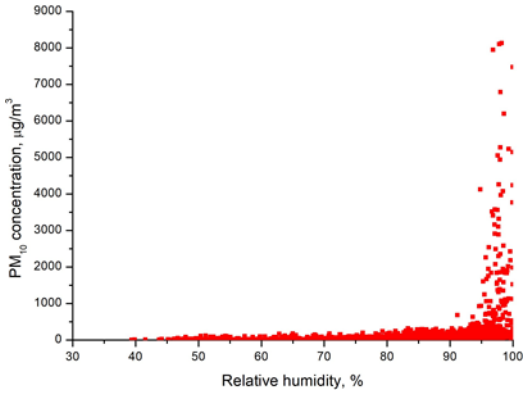


Figure 3: PM_{10} concentration vs relative humidity for the entire data set

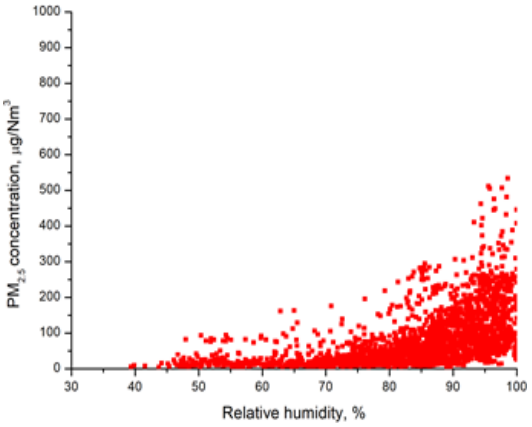


Figure 4: $PM_{2.5}$ concentration vs relative humidity for the entire data set

Considering the need to ensure the accuracy of measurements and the previous measurement results, it is necessary to determine a threshold value of relative humidity from which the measurements accuracy decrease. For this purpose, Pearson's correlation coefficients between relative humidity and PM concentrations were calculated for different data sub-sets. The results shown in Figure 5 indicate a stronger correlation for both types of PMs when relative humidity values are higher than 80%.

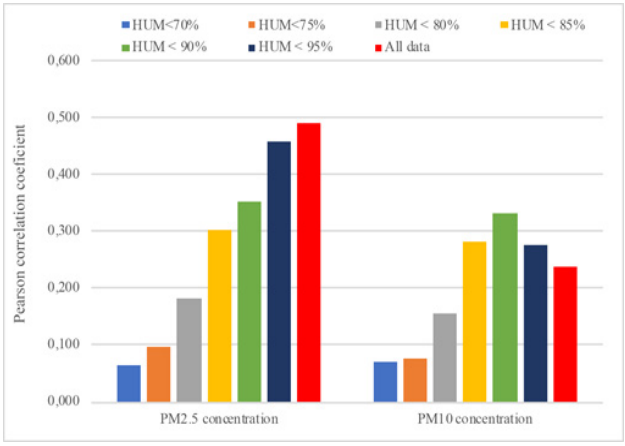


Figure 5: The absolute values of Pearson's correlation coefficients between $PM_{2.5}$ and humidity, respectively PM_{10} and humidity, for selected data sub-sets.

In order to perform a comparative analysis, the absolute value of the correlation coefficients was used, and the threshold value of relative humidity was established at 80%. The dataset that resulted consists of 591 values and represents 27.7% of the total values registered.

The variation of measured PM_{10} and $PM_{2.5}$ concentrations for the data sub-set corresponding to values of relative humidity lower than 80% are presented in Figure 6. For the same data sub-set, Figure 7 shows the PM_{10} and $PM_{2.5}$ concentrations function of relative humidity.

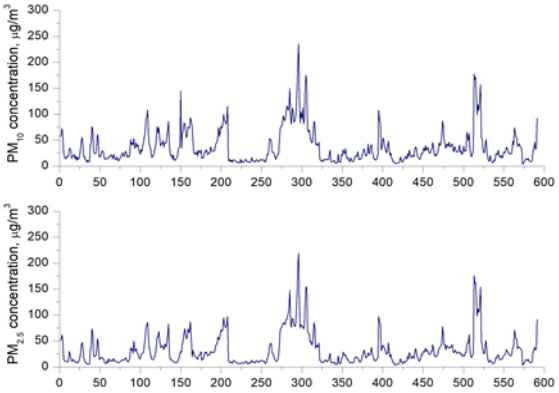


Figure 6: Variation of measured values for PM_{10} and $PM_{2.5}$ concentration for the data sub-set (humidity < 80%).

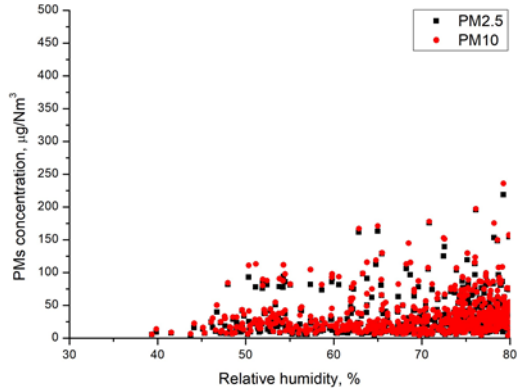


Figure 7: $PM_{2.5}$ and PM_{10} concentration vs relative humidity for the data sub-set (humidity < 80%).

Table 1: Correction factors for humidity classes

Class name	Range of variation, (%)	Correction factor
K1	≤ 80	1
K2	80-85	1.5152
K3	85-90	2.3008
K4	90-95	3.3807
K5	95-98	6.6515
K6	> 98	14.4549

In order to make corrections that eliminate the influence of humidity on the values of PM concentrations, it is proposed a division by humidity classes for which the values of correction factors have been estimated. The correction factor values for humidity classes were calculated as average values of the corresponding correction factors for the humidity values of each variation interval. The resulting values are presented in Table 1.

4.1. Algorithm for correcting concentrations of $PM_{2.5}$ and PM_{10}

Based on the information and data measured in the Tel-MonAer project, a correction algorithm (presented below) was developed for the concentration values of $PM_{2.5}$ and PM_{10}.

Step 1. **Determination of the correction factor for humidity**

Each measured value of the relative humidity falls into the corresponding humidity class (according to Table 1) and then the correction factor corresponding to the class is identified)

Step 2. **Correction for humidity of PMx concentration**

For each value of the **PMx** concentration measured, the following formula is applied:

$$Conc\ PMx\ corr\ H = \frac{conc\ PMx\ measured}{FC} \quad (1)$$

Where:

PMx – x fraction of particulate matter (e.g. $PM_{2.5}$ and PM_{10});

Conc PMx corr H – the value of PMx concentration as a function of humidity;

Conc PMx measured – the value of the measured PMx concentration;

FC – the value of the correlation factor.

Step 3. **Making the correction by reporting to the reference methods**

For each value of the concentration corrected in Step 2, the formula applies:

$$Conc\ PMx\ corr = f(Conc\ PMx\ corr\ H) \quad (2)$$

Where the function is specific to each type of sensor, pollutant and mediation period.

Step 4. **Calculation of the final concentration for the specified mediation interval.**

The average value of the corrected concentrations for the specified mediation periods (hour, day) is calculated.

For the application and testing of the calculation algorithm, the concentration data of $PM_{2.5}$, PM_{10} and relative humidity acquired using a Libelium SCP station (with OPC-N3 sensor) was used. The station was installed outside the building of the CAMPUS Center, within the Politehnica University of Bucharest (Figure 8). The data set used corresponds to the period March 13-May 13, 2019.

The corrected values of the concentrations of $PM_{2.5}$ and PM_{10} were calculated with the measurements made by the National Environmental Protection Agency, at the Morii Lake measuring point within the National Network for Air Quality Assessment (Figure 9).



Figure 8: Installation of the measuring equipment Libelium SCP at the CAMPUS center

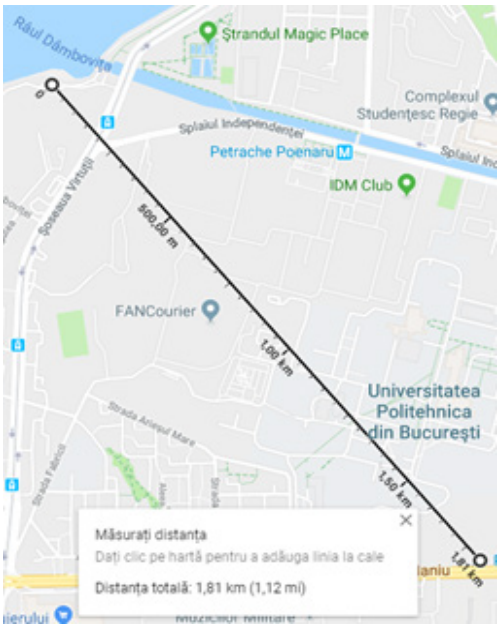


Figure 9: Distance between the Morii Lake monitoring point and the CAMPUS monitoring point

The correction algorithm was applied for $PM_{2.5}$ and PM_{10} and the results were compared with the values of the measured concentrations at Morii Lake measurement point in the National Air Quality Assessment network.

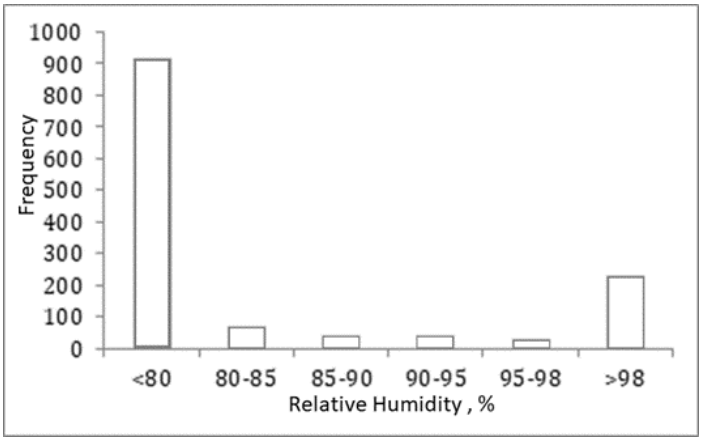


Figure 10: Histogram of relative humidity for the test period

4.2. Application of the algorithm for $PM_{2.5}$

During the analyzed period (March 13 – May 13, 2019), there were recorded hourly values of relative humidity (Figure 10) below 80% in 915 hours (69% of the total) and values greater than 98% in 231 hours (17.42%). Thus, the correction algorithm for humidity will lead to the modification of the values for 31% of the recorded values.

The hourly concentrations measured (Figure 11) during the testing phase of the algorithm recorded values below 7 µg / Nm^3 in 32.43% of hours, values less than 22 µg / Nm^3 being measured in 80.24% of the total number of hours. Also 43 values over 67 µg / Nm^3 were recorded.

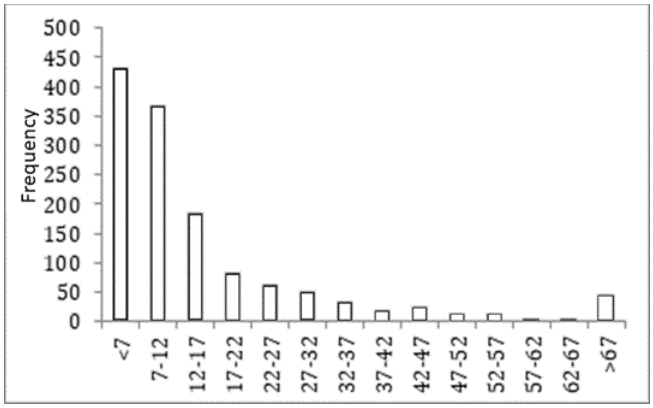


Figure 11: Histogram of the values of $PM_{2.5}$ concentrations measured during the test period

By applying the algorithm, the very small values (below 7 µg / Nm^3) of the measured concentrations were increased and represent 1.52% of the total. Values lower than 22 µg / Nm^3 represent 83.45% of the total number of hours. The number of values greater than 67 µg / Nm^3 was reduced to one value (Figure 12).

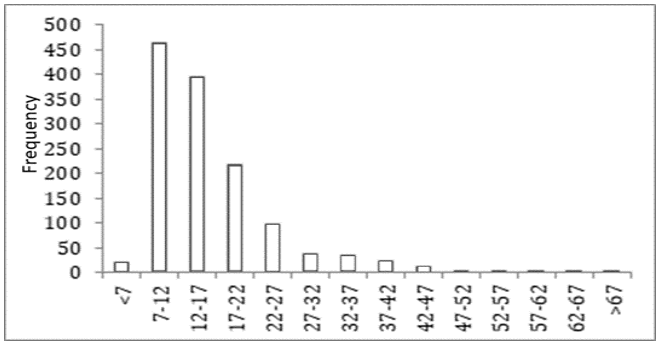


Figure 12: Histogram of $PM_{2.5}$ concentration values corrected after applying the algorithm

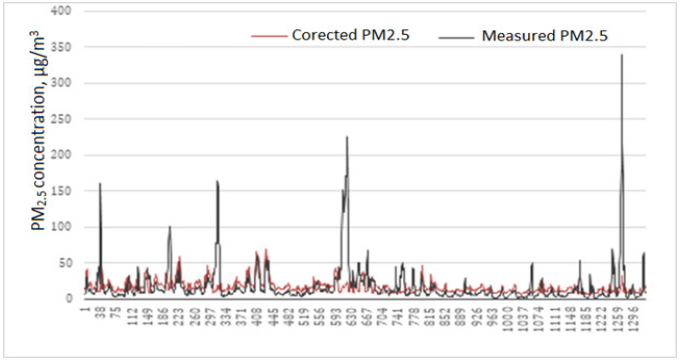


Figure 13: Measured and corrected concentration values of $PM_{2.5}$ for the test period

Figure 13 shows the values of $PM_{2.5}$ concentrations measured and corrected for the test period. It is observed the elimination of the extreme values generated by the increase of humidity and the increase of the small values which represents the elimination of the underestimation of the measured values.

4.3. Application of the algorithm for PM_{10}

The hourly concentrations measured (Figure 14) during the testing phase of the algorithm recorded values below 12 µg / Nm^3 in 44.49% of the hours, values less than 44 µg / Nm^3 being measured in 92.23% of the total number of hours. Also, 33 values of over 92 µg / Nm^3 were recorded.

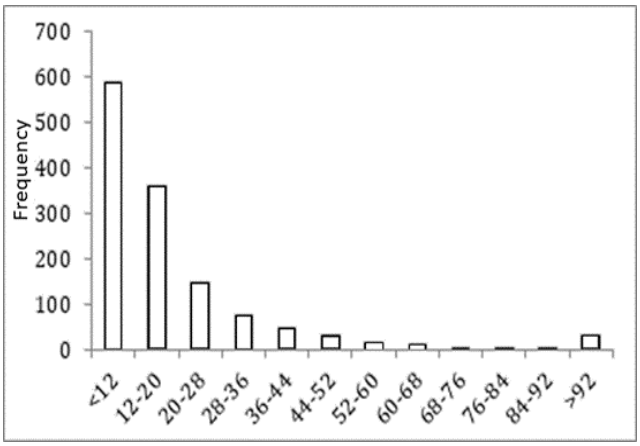


Figure 14: Histogram of PM_{10} concentration values measured during the test period

By applying the algorithm, the very small values (below 12 µg / Nm^3) of the measured concentrations were increased and represent 0.5% of the total. The number of concentrations greater than 92 µg / Nm^3 was reduced to a single value (Figure 15).

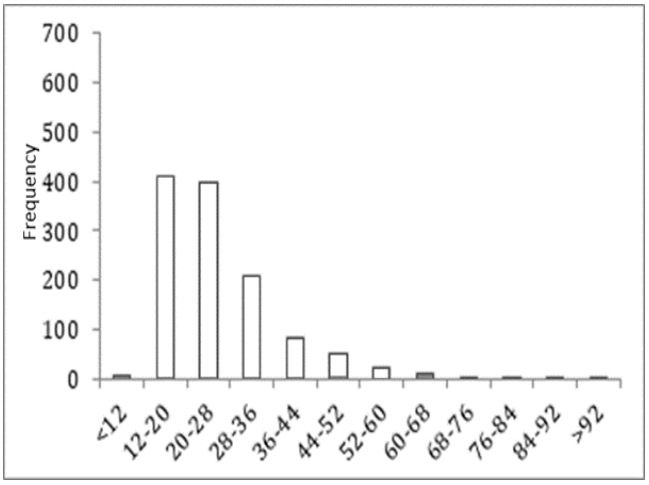


Figure 15: Histogram of PM_{10} concentration values corrected after applying the algorithm

The following figure shows the values of PM_{10} concentrations measured and corrected for the test period. It is observed the elimination of the extreme values generated by the increase of humidity and the increase of small values which represents the elimination of the underestimation of the measured values (Figure 16).

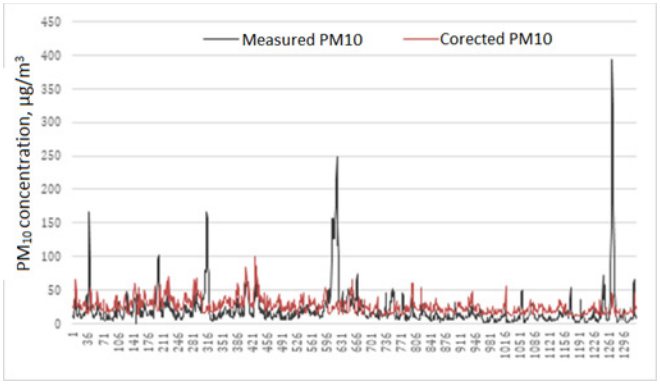


Figure 16: Values of concentrations measured and corrected by PM_{10} for the test period

4.4. Performance evaluation of the algorithm

The monitoring station at Morii Lake is urban-background type, the measured values being representative on an area with a radius of 1-5 km around the station. The CAMPUS Center where the Libelium sensors were located is within the representative area (1.8 km from the station). By placing it at a higher height, the effect of the pollution generated by car traffic was reduced, but it is also possible to reduce the measured values due to the height at which they were located.

The hourly $PM_{2.5}$ concentrations measured at the monitoring station at Morii Lake (Figure 17) were below the value of 22 µg / Nm^3 in 82.69% of the total number of hours.

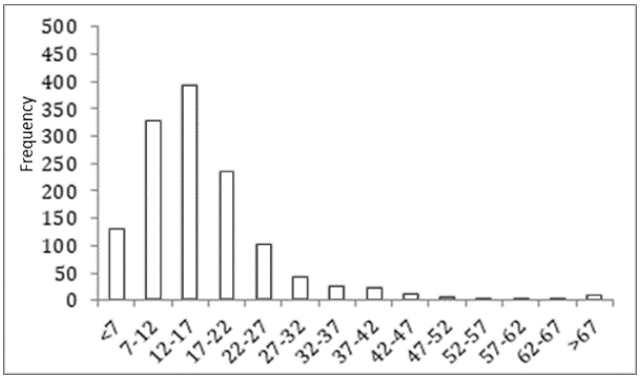


Figure 17: Histogram of the values of $PM_{2.5}$ concentrations measured at the Morii Lake point from the National Air Quality Network

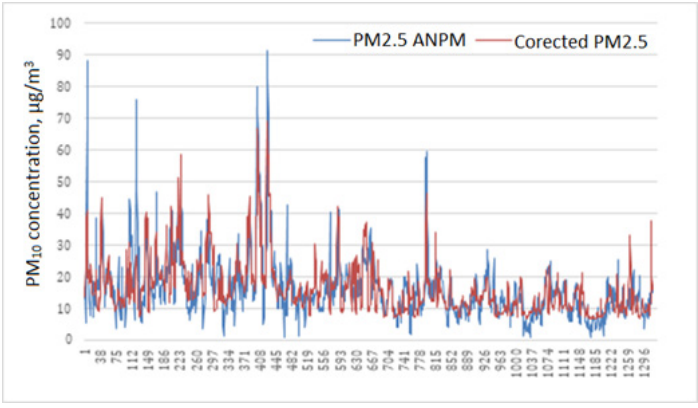


Figure 18: Values of corrected $PM_{2.5}$ concentrations and those measured by ANPM for the test period

The comparative graphical representation of the values of the corrected $PM_{2.5}$ concentrations and those measured by ANPM (Figure 18) during the testing period of the algorithm indicates close values and similar evolution trend.

The hourly PM_{10} concentrations measured at the Morii Lake monitoring station (Figure 19) were below the value of 44 µg / Nm^3 in 94.94% of the total number of hours.

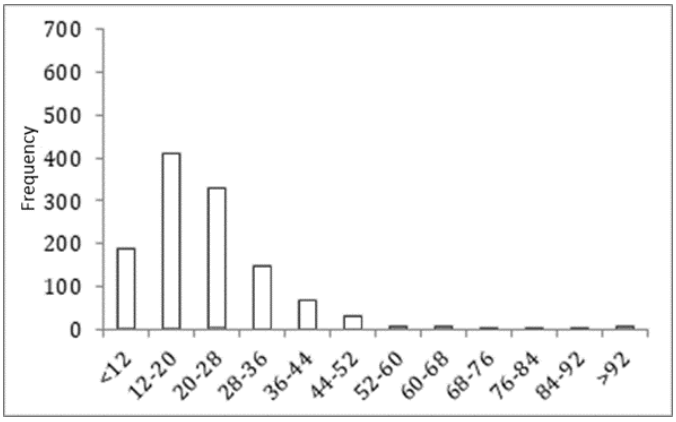


Figure 19: Histogram of the values of PM_{10} concentrations measured at the Morii Lake point from the National Air Quality network

The comparative graphical representation of the values of the corrected PM_{10} concentrations and those measured by ANPM (Figure 20) during the testing period of the algorithms indicates close values and similar evolution trend.

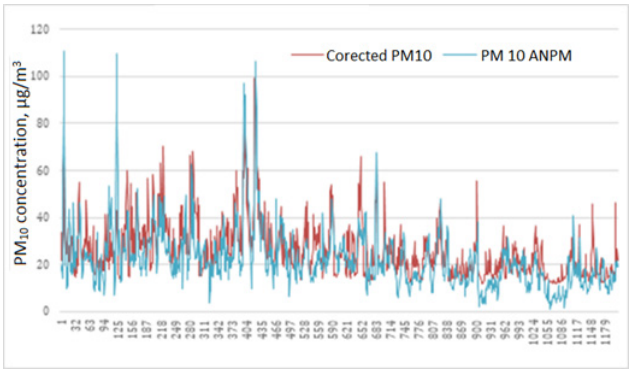


Figure 20: Values of corrected and officially measured concentrations of PM_{10} for the test period

The efficiency of the correction algorithm was evaluated at this stage by the value of the Pearson correlation coefficients (Table 2).

Table 2: Pearson's correlation coefficient values for the analyzed data sets

Concentration	Correlation coefficient	
	Measured values vs. ANPM values	Corrected values vs. ANPM values
$PM_{2.5}$	0.268	0.688
PM_{10}	0.216	0.663

This shows a 2.5-fold increase in the correlation coefficient for $PM_{2.5}$ concentrations, from a value of 0.268 (Libelium measured data vs. ANPM measured data).

For PM_{10} concentrations, the value of the correlation coefficient increased 3-fold, from 0.216 (Libelium measured data vs ANPM measured data) to 0.663 (Libelium corrected data vs. ANPM measured data).

5. Conclusions and Future Work

The influence of relative humidity on the PMs concentration values is significant for the devices based on laser measurement technology. For this type of IoT devices we propose a threshold value for relative humidity of 80% under which to consider PMs measured concentration as valid. For the situation when relative humidity has values over 80%, a specific algorithm was developed. The algorithm corrects the PMs measured values considering the type of the sensors and the value of humidity.

The correction algorithm was tested on a data set containing PMs concentration values and other meteorological parameters for a period of two months. The results show an increase of at least 2.5 times of the correlation coefficient between the corrected values and those measured by the reference station of the National Air Quality Monitoring Network.

Future work will consider further testing of the algorithm and the development of a specific ML algorithm for air quality predictions.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The work presented in this paper has been funded by Tel-MonAer project subsidiary contract no.1223/22.01.2019, from the NETIO project ID: P_40270, MySmis Code: 105976 and by the WINS@HI project, PN-III-P3-3.5-EUK-2017-02-0038.

References

[1] CSCS19: The 22nd International Conference on Control Systems and Computer Science, 28-30 May 2019, Faculty of Automatic Control and Computers, University Politehnica of Bucharest, Romania. https://cscs22.hpc.pub.ro/

[2] J. Ayres, "The mortality effects of long-term exposure to particulate air pollution in the United Kingdom," Report by the Committee on the Medical Effects of Air Pollutants, 2010.

[3] P. J. Landrigan, R. Fuller, N. J. Acosta, O. Adeyi, R. Arnold, A. B. Bald´e, R. Bertollini, S. Bose-O'Reilly, J. I. Boufford, P. N. Breysse et al., "The lancet commission on pollution and health," The Lancet, vol. 391, no. 10119, pp. 462–512, 2018.

[4] V. Hadjioannou, C. X. Mavromoustakis, G. Mastorakis, J. M. Batalla, I. Kopanakis, E. Perakakis, and S. Panagiotakis, "Security in smart grids and smart spaces for smooth IoT deployment in 5g," in Internet of Things (IoT) in 5G Mobile Technologies. Springer International Publishing, 2016, pp. 371–397. [Online]. Available: https://doi.org/10.1007%2F978-3-319-30913-2 16

[5] M. Ianculescu, A. Alexandru, and E. Tudora, "Opportunities brought by big data in providing silver digital patients with ICT-based services that support independent living and lifelong learning," in 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN). IEEE, jul 2017. [Online]. Available: https://doi.org/10.1109%2Ficufn.2017.7993817

[6] J. M. Batalla, K. Sienkiewicz, W. Latoszek, P. Krawiec, C. X. Mavromoustakis, and G. Mastorakis, "Validation of virtualization platforms for i-IoT purposes," The Journal of Supercomputing, vol. 74, no. 9, pp. 4227–4241, aug 2016. [Online]. Available: https://doi.org/10.1007%2Fs11227-016-1844-2

[7] J. S. Apte, K. P. Messier, S. Gani, M. Brauer, T. W. Kirchstetter, M. M. Lunden, J. D. Marshall, C. J. Portier, R. C. Vermeulen, and S. P. Hamburg, "High-resolution air pollution mapping with google street view cars:

exploiting big data," Environmental science & technology, vol. 51, no. 12, pp. 6999–7008, 2017.
[8] D. Fecht, A. L. Hansell, D. Morley, D. Dajnak, D. Vienneau, S. Beevers, M. B. Toledano, F. J. Kelly, H. R. Anderson, and J. Gulliver, "Spatial and temporal associations of road traffic noise and air pollution in london: Implications for epidemiological studies," Environment international, vol. 88, pp. 235–242, 2016.
[9] H.Lin,T.Liu,J.Xiao,W.Zeng,L.Guo,X.Li,Y.Xu,Y.Zhang,J.J. Chang, M. G. Vaughn et al., "Hourly peak pm 2.5 concentration associated with increased cardiovascular mortality in guangzhou, china," Journal of Exposure Science and Environmental Epidemiology, vol. 27, no. 3, p. 333, 2017.
[10] T. Watkins, "Draft roadmap for next generation air monitoring," Environmental Protection Agency, 2013.
[11] K.-H. Kim, E. Kabir, and S. Kabir, "A review on the human health impact of airborne particulate matter," Environment international, vol. 74, pp. 136–143, 2015.
[12] A. Y. Watson, R. R. Bates, D. Kennedy et al., Air pollution, the automobile, and public health. National Academies, 1988.
[13] S. Holgate, J. Grigg, R. Agius, J. R. Ashton, P. Cullinan, K. Exley, D. Fishwick, G. Fuller, N. Gokani, C. Griffiths et al., "Every breath we take: The lifelong impact of air pollution, report of a working party." Royal College of Physicians, 2016.
[14] F. Dominici, R. D. Peng, M. L. Bell, L. Pham, A. McDermott, S. L. Zeger, and J. M. Samet, "Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases," Jama, vol. 295, no. 10, pp. 1127–1134, 2006.
[15] J. G. Watson, J. C. Chow, H. Moosm¨uller, M. Green, and N. Frank, "Guidance for using continuous monitors in pm2. 5 monitoring networks," Nevada Univ. System, Desert Research Inst., Reno, NV (United States , Tech. Rep., 1998.
[16] B. Zhang, L. Jiao, G. Xu, S. Zhao, X. Tang, Y. Zhou, and C. Gong, "Influences of wind and precipitation on different-sized particulate matter concentrations (pm 2.5, pm 10, pm 2.5–10)," Meteorology and Atmospheric Physics, vol. 130, no. 3, pp. 383–392, 2018.
[17] K. Y. Ng and N. Awang, "Quantile regression for analysing pm10 concentrations in petaling jaya," Malaysian Journal of Fundamental and Applied Sciences, vol. 13, no. 2, 2017.
[18] C. Lou, H. Liu, Y. Li, Y. Peng, J. Wang, and L. Dai, "Relationships of relative humidity with pm 2.5 and pm 10 in the yangtze river delta, china," Environmental monitoring and assessment, vol. 189, no. 11, p. 582, 2017.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 183-190 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Configurable Process Model: Discovery Approach from Event Logs

Rabab Sikal[*,1], Hanae Sbai[2], Laila Kjiri[1]

[1]*AlQualsadi research team, ENSIAS, Mohammed V University of Rabat, Morocco*

[2] *FST, University Hassan II of Casablanca, Mohammedia, Morocco*

ARTICLE INFO

Article history:
Received: 30 September, 2019
Accepted: 29 December, 2019
Online: 22 January, 2020

Keywords:
Configurable Process Model
Process Discovery
Variability

ABSTRACT

In the domain of business process management, the configurable process model is widely used to optimize time and cost of business process models design, which is known as the concept of "reuse". Using process mining techniques for process model discovery helps to provide a better view on processes and improve quality of models. The majority of existing configurable model discovery approaches work intensively on control flow discovery as main process perspective without considering other perspectives such as resources and data, and do not propose a detailed discovery of variability elements. In addition, the configurable process model creation is generally done by merging variant models not directly from event logs, which is not the optimal way to get a reliable configurable process model. This paper presents an overview of new multi-perspective variability discovery approach. The approach respects the variability of different process perspectives and allows users to create a configurable process model directly from event logs.

1. Introduction

Process models are designed, managed, configured mainly through methods and tools provided by the Business Process Management (BPM) domain [1,2]. Actually, organizations are increasingly opting for event systems, also called process Aware Information Systems (PAIS) for the purpose of analyzing, supervising and optimizing the organization's processes [3]. Nowadays, rapid increase of business needs and fast changing of the enterprise environment derive the enterprise to face the challenge of saving time, reducing costs and minimizing errors of process management. Therefore, opting for reuse concept is a big requirement for making an optimal and flexible business process design [4,5]. Hence, the importance of taking into account the previous design experiences and do not design processes from scratch. In this sense, different approaches have proposed the reuse concept in process design while flexibility and adaptability are addressed in business process models [6, 7].

The configurable process model is defined as a single model that assembles all process variants in one model. It is also called "customizable process model" which means that this kind of process model regroups options, which can be configured by users to derive desired process variants. It represents commonalities and differences between all process variants, which offers flexibility and enables process design through reuse concept. The variation point is a configurable element of configurable process model. It represents where the variation occurs in the process model and represents all possible design choices. The configuration of the model consists of making choices of options for each configurable element according to specific requirements in order to derive individual and suitable model for the enterprise with minimal design effort. However, despite the diversity of approaches proposing creation and configuration of configurable process models, their management still requires a significant manual work in different steps (e.g. design, configuration and evolution).

Against this background, the techniques of process mining are introduced with the aim to automate the process management and minimize human intervention. The process mining uses data recorded in the event logs during the process execution in order to help organizations for discovering, checking conformance and enhancing their business processes.

Creating process model manually is a hard and redundant task, since the use of similar processes becomes more and more popular. For that reason, configurable process model discovery is used as an alternative for reusable process design. In this context, several

[*]Corresponding Author Rabab Sikal, rababsikal@gmail.com

works have proposed approaches for configurable process discovery [8, 9, 10]. However, existing works discover the control-flow as the main process perspective without considering other process perspectives like data and resources. In addition, the proposed variability discovery approaches do not present an explicit and detailed discovery of variability elements, namely variation points and variants. Moreover, existing approaches for configurable model creation use algorithms to construct configurable model by merging similar variant models, which result in large and complex models. In the light of these limitations, we have proposed in [11,12] a multi-perspective configurable process discovery approach with respect to variability of activities and resources. The comparative studies presented in these two works showed the lack of support for variability discovery for various process perspectives. This paper extends the work originally presented in the 2018 IEEE 5th International Congress on Information Science and Technology (CiSt) [11] that we complete by proposing an approach for configurable process model creation directly from event logs. The organization of this paper is as follows: section 2 defines the background of our research field, while section 3 reviews approaches working on configurable process model discovery. In section 4, we present our proposed configurable process discovery approach. Section 5 depicts our multi-perspective discovery framework and presents the algorithms for variability extraction and configurable model creation. Finally, Section 6 presents conclusions and some future directions.

2. Background

This section introduces three basic concepts used in this paper: process discovery, configurable process model, and variability. Then, we present briefly the four configurable process model discovery approaches described in [8].

2.1. Process discovery

Process mining is a set of techniques applied to extract data recorded in event logs. These data concern all process actions captured during process execution and are used to discover, monitor, and improve processes [1]. There are three main areas of process mining [1]:

- Process Discovery: the discovery algorithm takes an event log in input and produces a process model in output without using any additional knowledge.
- Conformance checking: verifies if an existing or discovered process model fits to its event log, or vice versa.
- Enhancement: extends and enrich existing process model, already discovered, by using information recorded in event logs.

In our approach, we focus on process discovery. There are two main kinds of process discovery:

Process discovery from one event log: it is the classical type of discovery. It allows for extracting one process model for each event log. However, this kind generates redundant processes [13,14].

Process discovery of a collection of event logs: this concerns the discovery of configurable process model. It requires firstly regrouping all event log that can belong to the same family and then applying techniques to discover the configurable process [8,9].

In Our work, we are interested in the configurable process model discovery from a collection of event logs.

2.2. Configurable process model & variability

The configurable process model represents shared/non configurable and unshared/configurable parts by all process variants in one global model. The configuration of configurable parts depends on the needs and the various constraints specific to the organization [15, 16]. Indeed, modeling all process variants and updating common process items cause redundancies and errors. Hence, the choice of configurable processes, generally presented by the merge of multiple process variants into a single process, is very useful to facilitate reuse and manage variability [17, 18]. Different extensions of process modeling languages have been developed for configurable process models representation, namely C-BPMN, C-EPC [15], C-YAWL [16] and configurable process tree.

The configuration consists of deriving individual process models corresponding to the different process variants from the configurable model. This operation is called individualization. The individualization is about blocking a given path of the model, so it cannot be taken or hiding activities they it can be skipped during the process execution. Designing a configurable process model consists in defining all the different variants of a given business process first and then integrating all of them into a single configurable model.

The variability is a key concept for configurable process model creation. It is represented by two main elements, namely, variation point and variants [13]. The variability defines and manages variable elements of business process [17]. Therefore, we define a process model that supports organizational, behavioral, functional and informational variability, as a multi-perspective configurable process model [1]. This can make the configurable process more explicit and valuable.

2.3. Approaches for configurable process model discovery

Process mining techniques for automated process discovery use data recorded in the event logs to represent the process behavior through a process model [6, 1]. Indeed, applying mining techniques for configurable process discovery is very useful, given the time saved and the effort reduced compared to conventional methods. In the literature, different approaches of configurable process discovery have been proposed. Buijs et al. [8] proposed four configurable process model discovery approaches, presented as follow:

- **Approach 1:** it is an approach initially proposed by Gottshalk [14]. The configurable process model discovered with this approach is the result of merging process models discovered from each event log.
- **Approach 2:** with the aim to improve the approach 1, the approach 2 merges all event logs and uses them to discover a common process model. Then, for every event log an

individual model is generated. Finally, the construction of the configurable model is done by merging the individual models.

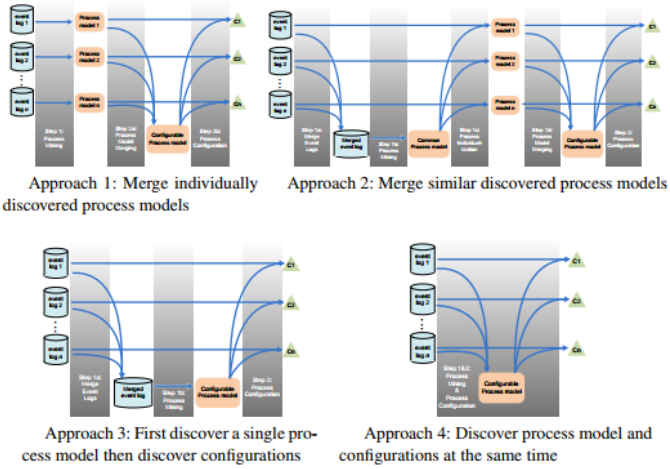


Figure 1: the four approaches of configurable process model discovery [8]

- **Approach 3:** it suggests the merging of different event logs into one merged event log. Then, the configurable is discovered. It captures the behavior of all model that describes the behavior of these event logs.
- **Approach 4:** it allows for the discovery of the process model and its configurations at the same time [8].

In this paper, our discovery approach belongs to the approach 4. It deals with redundancies and brings more flexibility in using discovery techniques to construct processes that capture variability.

3. Related works

In this section, we present several existing approaches for configurable process model discovery. Then, these approaches are evaluated according to four criteria. Finally, we discuss results and limitations.

3.1. Configurable process model discovery works

Several BPM studies are interested in using the paradigm of the "design by reuse" for the construction of a configurable process model. Some of them have proposed to construct the configurable process model by merging all process variants models into one model. Others have proposed to create a configurable process model using mining techniques on a collection of event logs. The author of [9] uses trace-clustering method for configurable process model discovery from collection of event logs. In [10], the author discovers configurable process fragments to avoid complex and large models. The author of [14] presents an approach using process mining and analysis techniques to merge two business process models into a single model for further process optimization. The approach in [16] merges the models of process variants to create configurable process model based on log files from various systems. The work provides suggestions for common and individual configurations. In [19], the author proposes two algorithms, one to compute merged models. The other, to extract digests from a merged model. The work [20] proposes an algorithm for constructing a configurable process by merging the process model of each variant, the process model generated by the algorithm is pre-annotated for the configuration step. The study of [21] splits the event logs in a cluster and for each cluster, A process model can be discovered. In case of large configurable process model, the model is reduced into a sub process model. Each sub-process model is configured independently to improve performance and to reduce complexity.

3.2. Comparative study

To summarize the previous section, the table 1 presents the principal points related to our approach and reached by every work. The approaches are evaluated according to four criteria defined as follows:

- **Variability discovery:** it indicates if the approach discovers explicitly the elements of variability (e.g. variation point, variants and variables).
- **Perspective discovery**: it presents the perspectives discovered by the approach. The main process perspectives are: control flow (C.F), resource (R), data (D) and configuration (C).
- **Configurable Model construction**: it indicates if the approach constructs configurable model.
- **Discovery approach for configurable model construction:** it indicates which approach, from the four approaches proposed by [8], is used for the construction of configurable model.

Table 1: comparative study

Works	Variability discovery	Perspective discovery				Configurable Process model construction	Discovery approach for configurable model construction			
		C.F	R	D	C		App 1	App 2	App 3	App 4
[9]		+	-	-	+	+	-	-	+	-
[10]	+	+	-	-	+	+	-	+	-	-
[14] [16] [19]	-	+	-	-	-	+	+	-	-	-
[20]	-	+	-	-	-	+	-	+	-	-
[21]	-	+	-	-	+	+	-	-	+	-

Table 1 is a summary of the evaluation criteria developed by each of the approaches presented in this section.

Variability discovery: The discovery of variability elements, namely variation point and variants, is present in few works. The works [10] discovers variability for configurable fragment of the process. The other studies [9, 14, 16, 19 20, 21] don't focus on variability elements in their discovery approaches.

Perspectives discovery: The studies [9, 10, 14, 16, 19, 20, 21] are limited to the discovery of control flow as a main process perspective, and discovery of its configurations. Thus, we notice the absence of support for discovery of other process perspective like resource and data.

Configurable process model construction: all the presented approaches propose the construction of configurable process model using different modeling language, like C-BPMN, C-EPC or process tree.

Discovery approach for configurable model construction: different works adopted different approaches for model construction. The works [14, 16, 19] construct configurable model based on approach 1. The [10, 20] adopt the approach 2 while [9,21] use approach 3.

3.3. Discussion

The analysis of the presented approaches shows that the majority of them do not present an explicit discovery of elements of variability and still limited to control flow discovery as a main perspective. While, other perspectives like resource still neglected and not integrated in the discovery approach. In addition, the construction of configurable process model is generally based on approach 1, 2 or 3 using merging techniques. We notice the absence of approaches for configurable model discovery based on the approach 4.

The limitations we conclude are as follow:

- The need for detailed discovery of variability elements, e.g. variation points, variation point types and variants. Whereas, discovering an explicit and detailed variability can be used to build configurable process models as well as its configurations. It can also be archived for potential process changes or improvements.
- Lack of multi-perspective discovery of process elements, namely data and resources. Instead of remaining focused on the analysis of the control flow, the extension of the configurable model with different process perspectives is of great importance. It helps analysts to manage the evolution of business process and to improve decision-making.
- The construction of the configurable model is based on merging individual models. Unlike, an approach that discovers configurable model directly from the event logs without merging individual models may provide a better model structure and better configuration options.

In the light of the presented limitations, we propose an approach with a detailed discovery of variability elements for different perspectives (control flow, resource). In order to enhance the variability discovery for other perspectives, we generate variability specification files for detailed variability. This can ensure traceability and optimize the process of changing or updating business process. In addition, our approach adopts the approach 4 proposed in [8] for optimal creation of configurable process model.

4. Preliminaries

In this section, we present formal definitions of basic concepts related to this work.

4.1. Event logs

Event logs are defined as files that store process data collected during process execution. The process mining techniques use data recorded in event logs for discovering process models, checking conformance between process model and its event log, detecting execution deviations or errors and observing social behaviors.

The table 2 illustrates an example of event log for a "purchase online" process. To buy an article the customer starts with "creating a personal account online (a)". After, the customer "choose products to buy (b)" and then "chose the payment method (c)" it can be "payment by card (d)", "payment by PayPal (e) " or "bitcoin payment (f)". Thereafter, the customer "confirms the payment (g)". If the payment is ok, "delivery service is activated (h)". If not, the customer must verify the payment data.

IS: Information System

Table 2: event log of purchase online process

Case ID	Activity	Resource	Date	Time	…
1	a	IS	23-05-2017	13:00:00	…
1	b	Jane	23-05-2017	13:10:00	…
2	a	IS	23-05-2017	14:00:00	…
1	c	Jane	25-05-2017	11:20:00	…
2	b	Ellen	25-05-2017	12:00:00	…
1	d	IS	25-05-2017	13:40:00	…
3	a	IS	25-05-2017	14:05:00	…
3	b	Adele	25-05-2017	15:00:00	…
1	g	Jane	26-05-2017	08:45:00	…
2	c	Ellen	26-05-2017	09:36:00	…
3	c	Adele	26-05-2017	11:00:00	…
2	e	IS	26-05-2017	11:15:00	…
3	f	Adele	26-05-2017	11:20:00	…
2	g	Ellen	26-05-2017	11:40:00	…
1	h	IS	27-05-2017	09:00:00	…
3	g	Adele	27-05-2017	09:30:00	…
3	h	IS	27-05-2017	10:00:00	…
2	h	IS	27-05-2017	11:20:00	

The data recorded in event logs present the execution history of one business process within an organization. A log case represents one process instance execution. The log represented in Table 2, records three different executions of the same process. Each process execution is called process instance, and referenced by an ID. The event log contains additional attributes, such as resource that executes the activity, date and time of the activity execution.

Definition 1: (Trace, Event log). Let A be a set of activities in some universe of activities. A trace σ ∈A∗ is a sequence of activities. An event log L∈T(A∗) is a multi-set of traces, i.e., an event log.

For instance, < a; b; c; d; g; h > is a trace that belongs to the event log in Table 2.

4.2. Log based relation

A log file is a set of traces. A trace can be defined as a sequence of events ordered chronologically and executed correctly. The execution order of activities in a process instance is of great importance. It helps to define dependency between activities and to capture all possible patterns encoded in the event log [1]. Based on the activities execution orders in traces, four ordering relations can derived from an event log: >L, →L, ||L and #L [22].

Definition 2: (Log-based ordering relations [22]) Let L ∈ M (A∗) be an event log over A. Let a; b∈A. The four ordering relations are precedence ($>_L$), causality ($\rightarrow_L$), parallelism ($||$) and choice ($\#_L$), defined as follows:

Precedence: $a >_L b$ if $\exists\sigma =< a1; a2; :::an >\in L$ and $1 \leq i \leq n - 1$ such that ai = a ^ ai+1 = b.

Causality: $a \rightarrow_L b$ if $a >_L b \wedge b >_L a$.

Parallelism: $a||_L b$ if $a >_L \wedge\ b >_L a$.

Choice: $a\#_L b$ if $a >_L b \wedge b >_L a$.

Based on the four log-based ordering relations, the discovered process model (figure 2) describes the behavior observed in the event log (table 2). The generation of the model is done by a discovery algorithm (e.g. [9,22]), and its representation by BPMN modeling language.

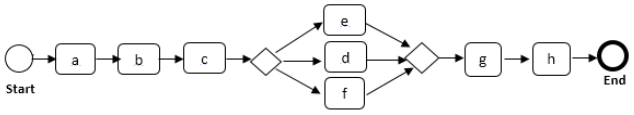


Figure 2: The process model corresponding to the event log in table 2

Event logs can be stored and exchanged using different forms of data source. MXML (Mining eXtensible Markup Language) is a standard notation for storing process attributes such as timestamps, resources and transaction types [23]. XES (eXtensible Event Stream) [24] is the MXML successor created to extend MXML.

4.3. Event log pre-processing

Logs are widely available in many applications, but the purpose of their creation and their level of details varies. To construct a configurable process model with meaningful behavioral patterns, the event logs must be pre-processed before using mining techniques.

- The elimination of confidential data is required before any data processing.
- The balance in the level of details in event logs is recommended. The generated process model has not to be highly detailed.
- The use of the same ontological concept for different sources of event logs.

5. Configurable process discovery approach

The construction of the process model can be done by merging models of process variants, which is complicated and error-prone especially when the number of variants is quite high. The approach we propose builds configurable process model from event log without merging exiting process models. It is based on event logs because of several reasons:

- Event logs are commonly available in Process Aware Information Systems (PAIS), such as : ERP, CRM and workflow management systems
- Business process models do not always exist. Therefore, techniques of merging cannot be applied.
- Event logs record process execution data exactly as it was executed in reality. The information recorded in the event logs is very useful for the business process design or configuration. For example, a priori process model does not present information like activity execution frequency, execution errors, and social behavior between users or services.

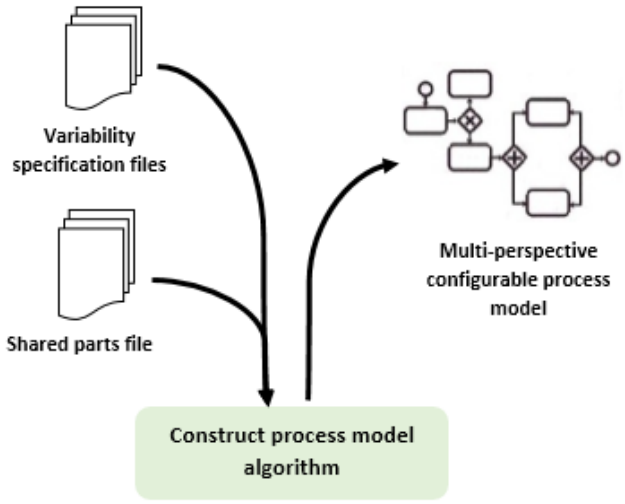


Figure 3: Construct configurable process model approach

In our approach (Figure 3), we discover configurable process model from event logs using two important files created directly from a collection of event logs: variability specification file and specification file of shared parts. The purpose behind the generation of these variability specification files is firstly to discover, in detailed manner, the variability of activity and resource and secondly to keep a record of variability for any forward configuration or update of the business process. The information recorded in the variability specification file represent variation points and variants of the configurable process while the specification file of shared parts represents the non-variable parts of the configurable process.

5.1. Framework architecture

The contribution presented in [11] proposes a variability discovery approach for business processes taking into account the

variability of resource perspective. In this paper, we shed light on our framework and its components. As well as, we describe the role of each component and we explain the interdependence between components. The discovery approach comes after, to create the configurable process model directly from event logs without using merging techniques.

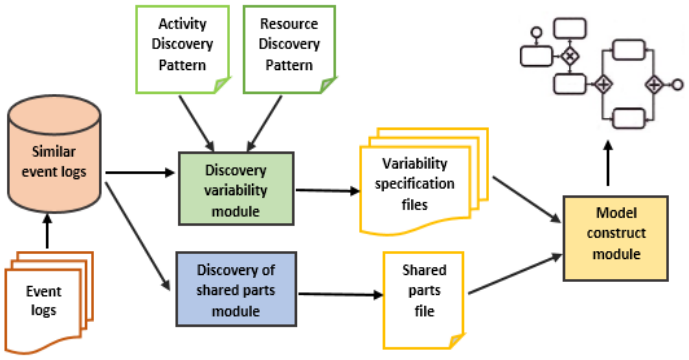


Figure 4: Architecture of the proposed framework [11]

The figure 4 presents the four components of the proposed framework:

- **Similar event logs:** It is a storage module of similar event logs collected and sorted by using existing techniques of clustering [9, 25, 26]. It takes as input a set of event logs and generates a set of pre-processed event logs.
- **Discovery variability module:** this module is important to discover the variability of different process perspectives namely activity, resource and data. It takes as input algorithms for activity and resource variability discovery and generates variability specification files for both activity and resource.
- **Discovery of shared parts module:** this module discovers the common parts shared by all process variants. It takes as input a set of similar event logs and generates specification file of shared parts of the process.
- **Model construction module:** this module constructs the configurable process model using the specification files of variability and shared parts.

The variability discovery module allows discovering the configurable fragments of the business process. It uses two algorithms to discover variability elements in detail. The first discovers the variability of activities [11] and the second discovers the variability of resources [23]. As output, the module generates i) a variability specification file, with variation points, variants, and variation point types for both activity and resource.

The shared parts discovery module discovers the common activities between all process variants. It generates ii) a specification file of common parts. This module allows discovering the non-configurable fragments of the business process.

Both variability discovery module and shared parts discovery module are essential for building the configurable model. The algorithm for configurable model creation uses i) and ii) for the configurable process model construction.

5.2. Discovery of variability

The discovery algorithm that we have introduced in [11] aims to provide an explicit variability discovery of activities. It discovers elements of variability such as variation points, type of variation point and variants. The algorithm uses discovery rules for each variability element [11]. These rules are used to define relations between activities and to construct the control flow process model.

Algorithm 1 : discovery of activity variability

```
1   Begin
2   Select the activity column
3   //Step 1: Discovery of variation point
4   For each activity do //From all activities of the process
5   For the direct activity successor do
6   Select successor1
7   // application of variation point discovery rule VP_rule
8   For each other_successor of the selected activity do
9     If successor1≠other_successor then
10              a variation point exists
11              save successor1
12              save other_successor
13    End if
14  //Step 2: discover of variants; application of var_rule
15   For each successor do
16    Define a variant
17  //Step3:name variation point as successor name
18  Name variation point=concatenation of all its
19  successors+vp as prefix
20  //Step4:define variation point type by application of
21  TVP_Parallel rule
22     If (successor1 is followed by other_successor)
23      and (other_successor is followed by successor1)
24     Then the point of variation is optional
25      End if
26     // application of TVP_Choice rule
27     If (successor1 is not followed by other_successor)
28     and (other_successor is not followed by successor1)
29     Then the point of variation is alternative
30     End if
31     End for
32  End for
33  End for
34  End
```

The discovery algorithm proceeds in different steps:

The algorithm discovers variation points based on the definition of activity successors from event logs. Then, it discovers the variants of variation points by defining the different successors identified in the previous step. After that, we concatenate the variants names to create the variation point name with add of the prefix “vp”. Finally, the algorithm discovers the type of variation point. For that, the log-based ordering relations [22] are used to discover relations between activities (choice or parallel).

The application of the discovery algorithm for activity variability generates variability specification file.

Definition 3: (variability specification file). Let A be a set of all activities of an event log L and a∈ A. Let VP, V be a two sets of finite activities, with VP ⊂ A, V ⊂ A and VP ⊔ V ⊆ A

Let o, a be the optional and alternative type of variation point as defined in [11], and let T= {t_o, t_a} be the type given to each variation point, respectively. Configurable fragment is defined as:

ConfigFrag(VP,T,V)= (a ∈VP, t∈T , (a_1,...a_n) ∈V n∈IN)

Variability specification file is a set of configurable fragments. It is defined as: ∑ ConfigFrag(VP,T,V)

This file contains details about variability elements of control-flow perspective. Different algorithms in our approach use this file for the discovery of resource variability and shared parts of the process variants.

5.3. Discovery of shared parts

To define shared parts of the process we implement algorithm based on the theory of sets in mathematics. It is defined as follow:

Let ET, EV and EC to be three sets of process activities. Where ET is the set of all process activities present in the event logs, EV is the set of activities present in the specification file of variability and EC the set common activities, we have: {E T = EV ⊔ EC } ⇔ {EV ⊂ ET , EC ⊂ ET , EV∩ EC =∅} ⇔{EC = ET\EV}

The algorithm uses the variability specification files and similar event logs. Which means that, the similarity between two activities name is maximum between their syntactic similarity and their linguistic similarity.

Algorithm 2 : Discovery of shared activities

```
1   init E(C) = ∅
2   begin
3       for each x ∈ E(T) do
4           if x not element of E(V) then
5           write x in E(C)
6           else continue
7           end if
8       end for
9     return E(C)
10  end
```

The algorithm parses the variability specification file of activities to discover the activities in common with all process variants, and then generates, as output, a specification file of common activities.

5.4. Creation of configurable process model

The discovery approach we propose constructs a configurable process model from event logs using the two files generated by the presented algorithm. The algorithm extracts the configurable fragment of the process model from the variability specification file of activities and extract the non-configurable fragment of the process model from the specification file of shared activities.

Algorithm 3 : Discovery of configurable model

```
1   Begin
2   Init Graph G
3   While E(C) ≠ ∅ do
4       for x ∈ E(C) do
5           if x ∈ E(V) then
6           G← create node x
7           define type of connector node
8           G ←create children nodes of  x successors
9           else
10          G ←create node
11          end if
12      end for
13    return graph G
14  end while
15  end
```

The algorithm parses the two generated specification files to construct the configurable model. It starts by selecting an element of shared parts and parses the variability specification file. If the element exists in the variability specification file, then the current element is a variation point, and the algorithm creates the corresponding node and its variants. If not, the activity is common of all process variants and the node is created.

6. Conclusion & future work

Many approaches were interested in control flow discovery, but few ones have been proposed for explicit variability discovery of different process perspectives. Given the importance of the variability for process reuse in business process management, the aim of our work is to propose a discovery approach of configurable process model from event logs, which provides a detailed information about variability elements in business process. In addition, the approach proposes a discovery algorithm for variability of other perspectives like activities and resources.

In the future work, we intend to work on the implementation of our approach with test results performed to evaluate its feasibility through different experiments. In addition, we will show the practical usefulness of our approach and publish a paper about the integration of resource variability in the discovered configurable process model.

References

[1] W.M.P. van der Aalst, Process mining: data science in action, 2nd edition Springer, 2016.

[2] M. Weske, Business Process process Management management concepts, languages, architectures. 2nd edition, Springer, 2005.

[3] M. Dumas, WM.P. Aalst, A.H.M. Hofstede, Process-Aware Information Systems Bridging People and software through process technology. John Wiley & sons, 2005.

[4] J. Becker, M. Kugeler, M. Rosemann, Process management: a guide for the design of business processes, Springer Berlin, Heidelberg, New York, 2003.

[5] Wil M. P. van der Aalst. "Workflow verification: Finding control-flow errors using petri-net-based techniques". In *Business Process Management, Models, Techniques, and Empirical Studies*, Springer-Verlag, 1806, 161-183, London, UK, 2000.

[6] R. Rychkova, S. Nurcan, "Towards adaptability and control for knowledge-intensive business processes: Declarative configurable process specifications", 1-10, IEEE Computer Society, 2011 https://doi.org/10.1109/HICSS.2011.452

[7] L. Seinturier, P. Merle, R Rouvoy, D. Romero, V. Schiavoni, J. B. Stefani. "A component-based middleware", 42(5), 559-583, May 2012.

[8] J.C.A.M. Buijs, B.F. van Dongen, W.M.P. van der Aalst., "Mining configurable process models from collections of event logs". In: Daniel F., Wang J., Weber B. (Ed.), Business Process Management. Lecture Notes in Computer Science, 8094, 33-48, Springer, Berlin, Heidelberg, 2013. https://doi.org/10.1007/978-3-642-40176-3_5

[9] Y.P.J.M. Van Oirschot, Using trace clustering for configurable process discovery explained by event log data, master thesis, University of technology Eindhoven, Netherlands, 2014.

[10] N. Assy, Automated support for configurable process models, PhD thesis, University Paris-SACLAY, France, 2015.

[11] R. Sikal, H. Sbai, L.Kjiri, "Configurable process mining: variability Discovery Approach", the 5th International Congress on Information Science and Technology,(IEEE CIST'18), Morocco, October 2018. https://doi.org/10.1109/CIST.2018.8596526

[12] R. Sikal, H. Sbai, L. Kjiri, "Promoting resource discovery in business process variability", proceedings of the 2nd international conference on Networking, Information System & Security, 48, 1-7. 2019. https://doi.org/10.1145/3320326.3320380
[13] F. Milani, M. Dumas, N. Ahmed, R. Matulevičius, "Decomposition driven consolidation of process models". In proceedings of the International Conference on Advanced Information Systems Engineering (CAiSE), 7908, 193-207. Springer, Berlin, Heidelberg, 2013.
[14] F. Gottshalk, W.M.P. van der Aalst, M.H. Jansen-Vullers, "Mining reference process models and their configurations". In R. Meersman, Z. Tari, P. Herrero (Ed.), On the Move to Meaningful Internet Systems: OTM 2008 Workshops, 5333, 263-272. Springer berlin Heidelberg, 2008.
[15] M. Rosamann, W.M.P. Van der Aalst, "A configurable reference modeling language". Information Systems, 32(1), 1-23, 2007. https://doi.org/10.1016/j.is.2005.05.003
[16] F. Gottshalk, W. M. P. van der Aalst, M. Jansen-Vullers, M. La Rosa, "Configurable Workflow Models", International Journal of Cooperative Information Systems, 17(2), 177-221. https://doi.org/10.1142/S0218843008001798
[17] H. Sbai, M. Fredj, L. Kjiri, "A pattern based methodology for evolution management in business process reuse". International Journal of Computer Science Issues, 11(1), 1-10, 2014.
[18] H. Sbai, L. El Faquih, M. Fredj, "A novel tool for configurable process evolution and service derivation". International Journal of Enterprise Information Systems (IJEIS), 15(2), 1-18, 2019.
[19] F. Gottshalk, W.M. P. van der Aalst, M. H. Jansen-Vullers, "Merging event-driven process chains". In R. Meersman, Z. Tari, P. Herrero (Ed.), On the Move to Meaningful Internet Systems: OTM 2008 Workshops, 5331, 418-426. Springer, berlin, Heidelberg, 2008. https://doi.org/10.1007/978-3-540-88871-0_28
[20] W. Derguech, S. Bhiri, E. Curry, "Designing business capability-aware configurable process models". Information Systems, 72, 77-94, 2017. https://doi.org/10.1016/j.is.2017.10.001
[21] M. Arriagada-Benítez, M. Sepúlveda, J. Munoz-Gama, J.C. A. M. Buijs, "Strategies to automatically derive a process model from a configurable process model based on event data". Applied sciences, 2017. https://doi.org/10.3390/app7101023
[22] W.M.P. van der Aalst, T. Weijters, L. Maruster. Workflow mining: Discovering process models from event logs. *16 (9)*. IEEE Transactions on Knowledge and Data Engineering, 2004. https://doi.org/10.1109/TKDE.2004.47
[23] B. F. Van Dongen. "A meta model for process mining data". In *Proceedings of the CAiSE WORKSHOPS*, 309-320, 2005.
[24] H.M.W. Verbeek, J.C.A.M. Buijs, B.F. van Dongen, W.M.P. van der Aalst. "XES, XESame, and ProM 6". In P. Soffer and E. Proper, editors, *Information Systems Evolution*, vol.72 of *Lecture Notes in Business Information Processing*, 60-75. Springer-Verlag, Berlin, 2010.
[25] M. De Leoni, W.M.P. van der Aalst, "A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs". In Information Systems, 56, 235-257. Oxford, UK, 2015. https://doi.org/10.1016/j.is.2015.07.003
[26] A. Appice, D. Malerba, A co-training strategy for multiple view clustering in process mining. IEEE Transaction on Services Computing, 9, 832-845, 2016. https://doi.org/10.1109/TSC.2015.2430327

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 188-196 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Hybrid Solar Thermal/Electricity Automated Oven

Ajah Onu Victor[*,1], Emenike Ejiogu[2]

[1]*Department of Electrical Engineering, University of Nigeria Nsukka, Nigeria*

[2]*African Centre of Excellence for Sustainable Power & Energy Development, University of Nigeria Nsukka, Nigeria*

ARTICLE INFO

Article history:
Received: 14 January, 2020
Accepted: 28 February, 2020
Online: 11 March, 2020

Keywords:
Sliding mode control
Solar thermal collector
Control law

ABSTRACT

This paper presents hybrid solar thermal/electricity automated oven. The work compares sliding mode control (SMC) to traditional PID control of the oven system using MATLAB/Simulink 2014b model. SMC control method shows faster rise and settling time. The control technique has been designed to automate change of temperature level of the oven by accepting multiple reference inputs. This has been implemented using microcontroller programed in C++. Flat-plate solar thermal collector and paint curing oven have also been implemented for experimental test. The results obtain from the experiments are in conformity to the simulation results. The collector delivers 43.27% of the total energy required to operate the oven up to 120°C at 200W/m2 solar insolation. Calculation of payback period for commercial viability on the cost of the collector installation indicates a year and four months for an average insolation of seven hours per day. These results demonstrate that the research work is effective and solves the problem of temperature control in curing modern organic paints and the challenge of power requirements in the operation of oven.

1. Introduction

This paper is an extension of work originally presented in 2019 IEEE International Conference on Sustainable Energy Technologies and Systems (ICSETS) [1].

Precise temperature control of ovens is very essential to industries and firms that use them for productions. Paint curing for instance involves converting applied wet or powdered paint to dry and hard film. Paint may cure by solvent loss, chemical reaction, oxidation, melting and re-solidifying, or melting and crosslinking [2]. It may take few hours for liquid paints to dry but it takes days to months to cure properly. To reduce cure time, paint curing ovens are used to speed up the process under elevated temperature, and energy is usually required for reliable and stable curing of industrial coatings [2, 3].

The primary challenges of curing and drying ovens generally are temperature control and energy optimization. Various temperature control methods such as traditional PID control system, Adaptive and Fuzzy Algorithm, fuzzy-PID control system, etc. [4, 5, 6] have been proposed in other works. Also, there are various methods energy can be made available to raise the temperature of oven and its content to required degree [7, 8]. The results obtained from the applications of these proposals show slow response time, delayed settling time, high overshoots and oscillations outside prescribed tolerance limits. Hence, they have not adequately met up with constraints in modern coatings and the complexity of curing ovens, especially when the different subsystems - heaters, air circulating fans, temperature sensors, etc. - have to work together to control temperature.

Moreover, as paint curing processes continue to improve, the tolerance limits have become much tighter from ±14°C in liquid-based paints to ±5.6°C in organic and powder paints [3, 9] which demands more robust control method that can track the set-point more accurately. On the other hand, Sliding Mode Control (SMC), a robust non-linear control method, has recently found wide applications to automatic control problems; largely due to its simple algorithm, order reduction, decoupling design procedure, disturbance rejection, insensitivity to parameter variations, simple implementation, high reliability and fast response [10, 11, 12].

The application of SMC to oven temperature control would provide the needed remedy since its comparison to the earlier mentioned control methods in other areas proved that SMC gives

[*]Ajah Onu Victor, University of Nigeria Nsukka, Address, Contact No: +2348036878453 & Email: ajahonuvictor@gmail.com

improved system performance [13, 14]. This work applies sliding mode control to paint curing oven temperature in other to overcome the inefficiencies of existing control methods. Sliding mode control (SMC), has many advantages and can track the desired temperature within limits. Solar thermal collector is incorporated to electric heating system for energy optimization.

2. Modeling and Simulation

2.1. Oven Modeling

Figure 1 is the block diagram model of energy flow in a convective oven system. The following assumptions were made in the modelling [15]:

i. That all objects in the chamber are under thermal equilibrium condition.
ii. The materials that compose the oven and the work-piece have constant physical properties and do not depend on temperature.
iii. The heat capacity $C = c_p \times m$ of insulators and supports of the oven are considered in the balance energy equation, where c_p and m are the specific heat and the mass of each component in the oven respectively.
iv. The change of temperature T with respect to time, t has the following requirements:
   a) $\frac{\mathrm{dT}}{\mathrm{dt}}$ is maximum at t = 0, it is a non-increasing function of t and
   b) $\frac{\mathrm{dT}}{\mathrm{dt}} \to 0$ when $t \to \infty$.
v. The energy of the chemical or phase transformation is neglect. It is assumed that no work enters or goes out from the oven

a) It is considered that the cavity and both air and body surfaces place in it are in thermal equilibrium.
b) The physical properties, as thermal conductivity and specific heat of oven components and work-piece materials, are constant.
c) The temperature change in the inner part of the oven satisfies the following restrictions:
   i. $\frac{\mathrm{dT}}{\mathrm{dt}} = \varepsilon \geq 0$.
   ii. ε is small in the heating process.
   iii. $\varepsilon(t)$ is maximum at t = 0, i.e., it is a decreasing function of t.
   iv. $\varepsilon(t) \to 0$ When $t \to \infty$.

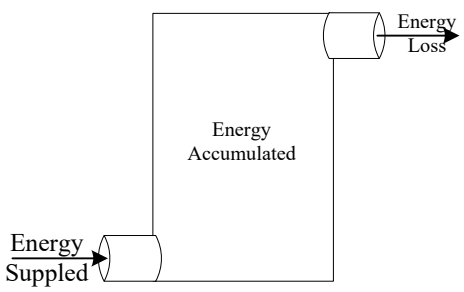


Figure 1: Convective oven energy flow model

The energy supplied E_s, to the convective oven is the sum of the accumulated energy E_u, and the energy loss E_L.

$$E_s = E_u + E_L \quad (1)$$

$$E_u = E_s - E_L = C_T \frac{dT}{dt} \quad (2)$$

The maximum heat energy loss

$$E_L = \frac{T - T_a}{R_T} \quad (3)$$

where,

E_s, E_u *and* E_L *are in Joules*

$C_T = heat\ capacity = c_p \times m\ (of\ insulators, support,)$

$c_p = specific\ heat\ capacity\ in\ (kJ/kgK)$

$m = mass\ (in\ kg)$

$T =\ temperature\ in\ the\ oven\ (in\ °K)$

$T_a = temperature\ of\ the\ surrounding\ (in\ °K)$

$R_T = total\ thermal\ resistance\ of\ the\ oven\ \left(in\ \mathrm{K}\cdot\frac{m^2}{\mathrm{W}}\right)$

Putting (2) & (3) into (1) gives (4)

$$E_s - \frac{T - T_a}{R_T} = C_T \frac{dT}{dt} \quad (4)$$

Re-arranging (4) gives (5) which is the temperature evolution of the oven.

$$R_T C_T \frac{dT}{dt} + T\ = E_s R_T + T_a \quad (5)$$

The maximum oven temperature and time constant are expressed by (6) and (7) respectively.

$$E_s R_T + T_a = T_{max} \quad (6)$$

$$R_T C_T = \tau \quad (7)$$

Thus, (5) becomes (8).

$$\tau \frac{dT}{dt} + T = T_{max} \quad (8)$$

Since τ & T_{max} are constants, (8) is a first order ordinary differential equation of the oven system and its solution is shown in equation (9).

$$T(t)\ = T_{max}\left(1 - \exp\left[\frac{-t}{\tau}\right]\right) + T_a \exp\left[\frac{-t}{\tau}\right] \quad (9)$$

Putting (6) into (9) gives (10).

$$T(t) - T_a = E_s R_T \left(1 - \exp\left[\frac{-t}{\tau}\right]\right) \quad (10)$$

Equation (11) is the maximum temperature rise, hence the temperature rise of the system at any time instant is obtained by (12).

$$E_s R_T = \theta_{max} \quad (11)$$

$$\theta(t) = \theta_{max}\left(1 - \exp\left[\frac{-t}{\tau}\right]\right) \quad (12)$$

Taking the Laplace transform of equation (8) gives (13) which is the transfer function of the system.

$$\frac{T_{(S)}}{E_{(S)}} = \frac{R_T}{\tau s+1} \quad (13)$$

2.2. Controller Design based on Sliding Mode Control

Equation (14) defines the sliding surface for nth order system [16, 17].

$$S = \left(\frac{d}{dx} + \lambda\right)^n \times e \quad (14)$$

Where,

$S = sliding\ surface$

$\lambda = a\ positive\ integer$

$n = the\ order\ of\ the\ system$

$e = error$

Re-arranging (5) gives (15), a first order equation.

$$\frac{dT}{dt} = \frac{1}{\tau}\{-T + E_s R_T + T_a\} \quad (15)$$

T is the actual output, taking T_d as the reference input, the error,

$$e = T_d - T \quad (16)$$

Substituting for n and e into (14) gives (17).

$$S = \left(\frac{d}{dx} + \lambda\right)^{1-1} \times (T_d - T) \quad (17)$$

$$\therefore S = 1 \times (T_d - T = 0) \quad (18)$$

The equivalent control of SMC is obtained by taking the derivative of the sliding surface as shown in (19).

$$\dot{S} = \dot{T}_d - \dot{T} = 0 \quad (19)$$

T_d is a constant $\therefore \dot{T}_d = 0$

$$\dot{S} = -\dot{T} = 0 \quad (20)$$

Putting (15) into (20) gives (21)

$$\dot{S} = \frac{1}{\tau}\{T - R_T E_s - T_a\} = 0 \quad (21)$$

Making E_s the subject in (21) gives the equivalent control of the system (22).

$$E_{s-equ} = \frac{C_T}{\tau}\{T - T_a\} \quad (22)$$

The switching control of SMC is given by (23).

$$E_{sw} = \begin{cases} Msign(S) & S > 0 \\ -Msign(S) & S < 0 \end{cases} \quad (23)$$

Where,

S is the sliding surface and

M is calculated based on Lipschitz function [18, 15].

SMC control law is given by (24).

$$U(t) = U_{equ} + U_{sw} \quad (24)$$

The control law in (25) is the summation of the E_{s-equ} and the switching control E_{sw} divided by the coefficient of E_s in (21)

$$E_s(t) = \frac{C_T}{\tau}\{T - T_a\} + \frac{\tau}{R_T}\left(Msign(s)\right) \quad (25)$$

The stability of the system is verified using Lyapunov [18]

Taking the sliding surface S, as the Lyapunov function candidate $\frac{1}{2}S^2$,

It will evaluate to common stability values:

$\frac{1}{2}S^2 > 0$ the system is unstable

$\frac{1}{2}S^2 = 0$ the system is stable

$\frac{1}{2}S^2 < 0$ the system is asymptotically stable

$\frac{1}{2}S^2 = \dot{S}S \Longrightarrow \frac{1}{\tau}(T - E_s R_T - T_a)(T_d - T)$

$\frac{1}{\tau}(T * T_d - T_d * E_s * R_T - T_d * T_a - T^2 + T * E_s * R_T + T * T_a)$

Since the highest power of the function $-T^2$ is negative, $\frac{1}{2}S^2 < 0$; therefore, the system is stable.

2.3. Simulation using MatLab/Simulink

Simulation of the oven system carried out in MATLAB/Simulink, as shown in the Simulink model description in Figure 2. The figure shows the reference temperature T_d of 120℃ as the desired set-point. The reference is compared to the actual oven temperature from the feedback loop to generate error signal. The error e, is used by the sliding mode controller to generate commands that regulates the quantity of heat energy that is delivered to the oven chamber.

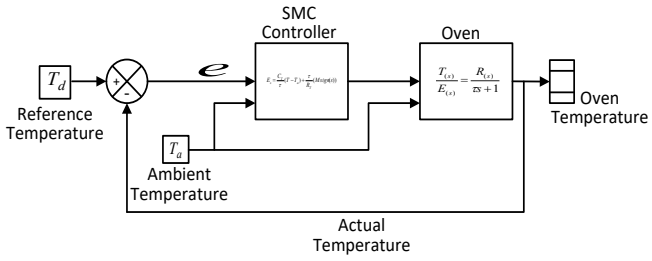


Figure 2: Simulink Model of SMC controlled Oven

The SMC controller is a Simulink model sub-system designed using (25). The ambient temperature, T_a is modelled in (26).

$$A_1 + A_2 \sin\varphi \quad (26)$$

Where,

$A_1 = 25℃$ (Average room temperature)

$A_2 = 8℃$ (Chosen amplitude of variation)

$\varphi \leq 0 \leq 360$ (In degrees)

The oven block is a Simulink model sub-system too. It is created using (13) which is the transfer function describing the oven system behavior. The oven temperature block is a scope that logs the temperature evolution in the oven.

The simulation is repeated after the SMC controller sub-system has been replaced with Simulink PID block for comparison.

2.4. Simulation Results

The simulation results for the SMC controller is shown as a plot of temperature against time in Figure 3. From the graph the controller drives the system temperature to rise continually until the reference temperature of 120℃ is reached. At steady-state the characteristic chattering effects of sliding mode controllers is seen as reviewed by a zoom-in of the plot. The highest and lowest peak of the chattering is ±0.3℃ which are within the tolerance limits of most of the sensitive materials that are baked in oven such as some organic paints that require ±5.6℃.

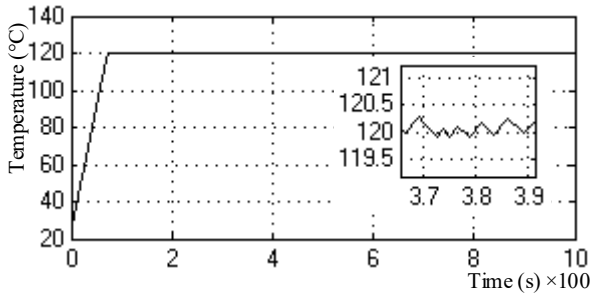


Figure 3: Plot of temperature vs time for SMC controlled oven

However, the simulation results after the SMC controller has been replaced by PID block shows high amplitude of oscillation at the steady-state region. The plot of temperature-time for the simulation is shown in figure 4.

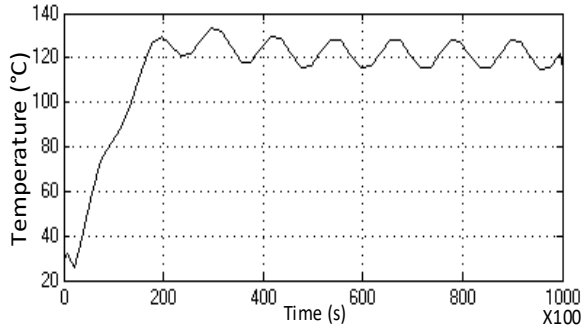


Figure 4: Plot of temperature vs time for PID controlled oven

The rise time for the PID controller is very long (11700s as against <100s for the SMC.

Table 1 shows comparison of further analysis results generated from Matlab tools like stepinfo command.

Table 1: Comparison of SMC and PID controller for oven

Response	SMC Value	PID Value
Rise Time:	98.39 (s)	11700 (s)
Settling Time	124.14 (s)	64800 (s)
Settling Min	119.709°C	115.026
Settling Max	121.144°C	133.36°C
Overshoot:	0.9535%	11.13%
Undershoot:	0.2425%	4.145%
Peak vale:	121.144°C	133.36°C

Some important control factors have been compared as reviewed in Table 1.

i. Rise time – this is the time taken for the system to rise from 10% of the steady-state value to 90% of the steady-state value (the steady-state value is 120℃). From the analysis the SMC controlled oven has 98.39 s rise time while the PID controlled oven has a rise time of 11700 s.

ii. Settling time – this is the taken for the system to reach and remain within tolerance limits of steady-state value. The SMC controlled system has 124.14s and the PID controlled system has 64800 s to settle.

iii. Settling max and min – are the highest and lowest temperature values in the oven after the system has reached steady-state.

iv. Overshoot and undershoot – are the differences between the upper and lower limits of the system response and steady-state value expressed in percentage.

v. Peak value – is the maximum temperature value attained inside the oven chamber.

The analysis results are evidence that the SMC controller performs better than PID controller for oven system. Hence, the system implementation is based on sliding mode control.

3. Implementations

3.1. Software Designs

The simulation results of sliding mode control of oven temperature in MATLAB/Simulink 2014b indicates that the control method is an improvement on the existing control methods in for oven temperature control. Three key implementations that have been designed and built to verify SMC oven control on a practical oven system; these are convective paint curing oven, flat-plate air solar thermal collector and sliding mode controller unit.

Figure 5 and Figure 6 are the screen captures of the AutoCAD drawings of the oven and the collector respectively. The dimensions of the oven are a height of 1200mm, a length of 900mm and breadth of 600mm. It also has 1000mm length and 400mm breadth door framing; exhaust air channel of 100mm x 100mm x 1000mm and insulation layer thickness of 50mm.

Figure 5 shows control box that houses the controller, the keypad and display unit. Other things visible from the drawing are the power distribution box, the oven door lock and heater access door.

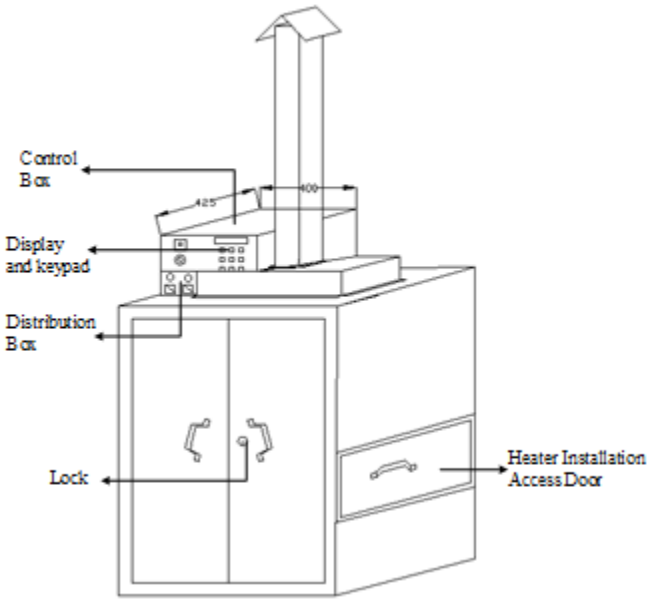


Figure 5: AutoCAD drawing of Convective paint curing oven

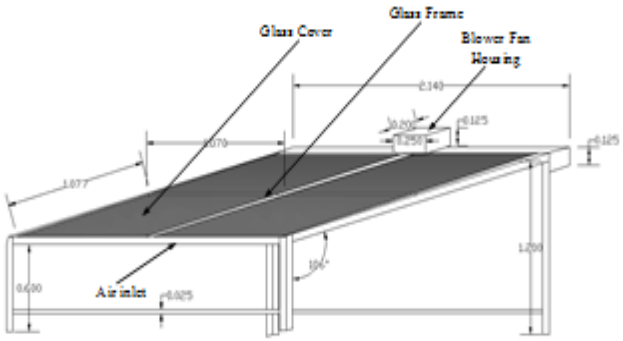


Figure 6: AutoCAD drawing of lat-plate air solar thermal collector

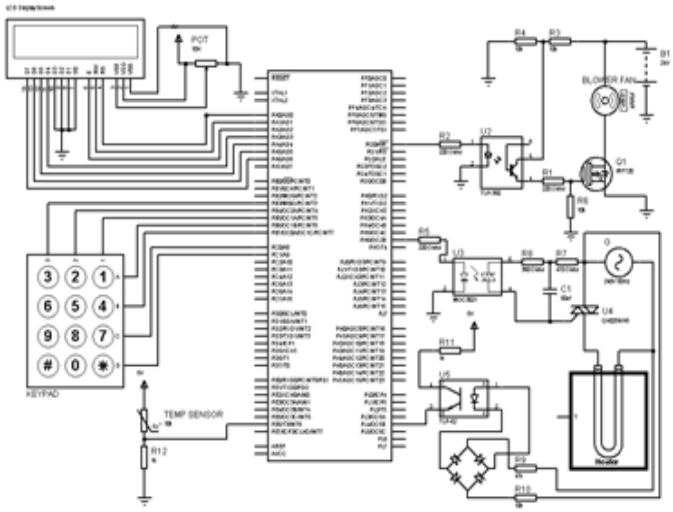


Figure 7: Proteus design of controller circuit diagram

Figure 6 shows 2140mm x 2140mm flat-plate collector tilted at 16° angle for higher reception of radiation since extraction is by air forced. The drawing shows the glass cover, glass frame, extraction fan housing and inlet air channel.

The design software is also used to generate the bill of materials (BOM), require for the implementation.

Figure 7 is screen capture of the control circuit diagram as designed in proteus software. It shows the microcontroller which is programmed using c++, a 16x2 LCD display, 3x4 matrix keypad electric heater control circuit and extraction fan controller for solar thermal collection.

3.2. Practical Implementation

The BOM generated from software is used to purchase the required materials need for the practical implementation. Figure 8(a) is a picture of the setup of the complete oven system with solar thermal collector and the control box. In the figure the chimney of the exhaust channel of can be seen on top of the oven and the control panel with the power distribution box mounted on the top left corner. The collector is place side by side with the oven and insulated host used to feed hot air extracted from the collector into the oven.

A zoom-in of the control box section is shown in Figure 8(b). It reviews the display, control box power switch, reset button, keypad and USB programmable port. The main power switches and fuse boxes are equally visible below the control box.

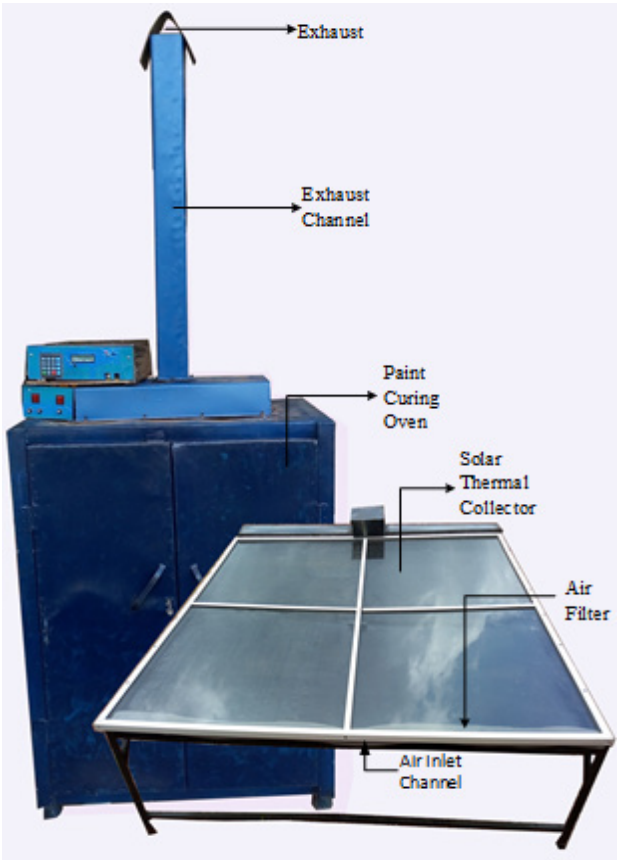


Figure 8(a): Picture of implemented system setup

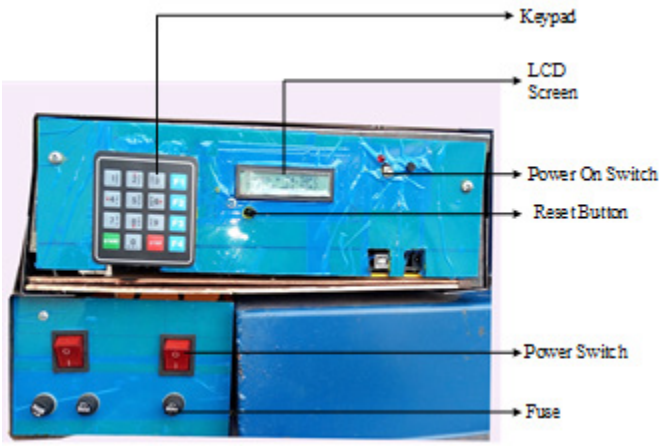


Figure 8(b): Picture of control box and power distribution box

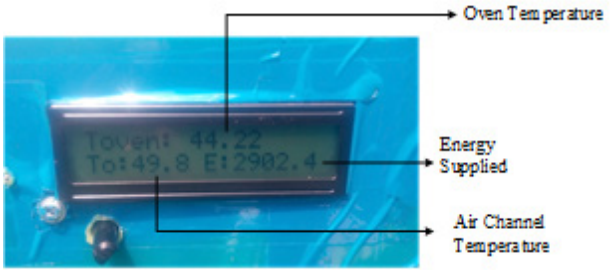


Figure 8(c): Picture of zoom-in display unit

Further zoom-in shown in Figure 8(c) is the real-time readings in the oven chamber as displayed by the LCD screen. They are the values obtained while the oven is powered by the solar collector alone. The value on top, Toven, is the temperature on the oven at the instant of taking the snapshot, the value at the lower left, To, is the temperature of the hot air extracted from the collector and the value at the lower right of the screen, E, is the total energy delivered by the collector.

3.3. Program Flow of the Controller – Automated Oven

The control implementation allows for setting two temperature levels of temperature reference points which the oven will automatically migrate from one level to another. The flowchart algorithm in Figure 9(a) shows the program flow. It starts by initializing the variables, T1, T2, t1 and t2. T1 is the first stage reference temperature for the cure (in Celsius) of paint that requires multiple temperature levels for proper cure. T2 is the second stage set-point; t1 is the time require for the paint to soak in the first temperature point; t2 is the soak time at the second temperature level. At the completion of initialization the processor reads the keypad and displays the numeric entry to the LCD screen. The processor evaluates the oven power control pin and toggle if need. This operation turns ON the air circulating fan in the oven at low speed. The processor reads the temperature sensors, which are negative temperature coefficient thermistors chosen because of their high sensitivities, fast responses, accuracies, and because they are suitable for temperature range of 20°C to 120°C of the paint cure oven built. The sensors data read are logged to external storage for analysis and simultaneously used in the sliding mode control subroutine. The display is also updated in real-time. The SMC subroutine generates the command signals the processor uses regulate the speed of the extraction fan, circulating fan and the electric heater. The processor compares actual oven temperature to the first reference temperature, T1 of the oven; once the reference is reached a timer is set and the SMC maintains the actual temperature in the oven at that reference point until the set time is elapsed. At the expiration of the timer, the screen is displays "First Stage Completed" text and the SMC subroutine is updated with the second reference point which will enable the oven to automatically maintain a new temperature profile.

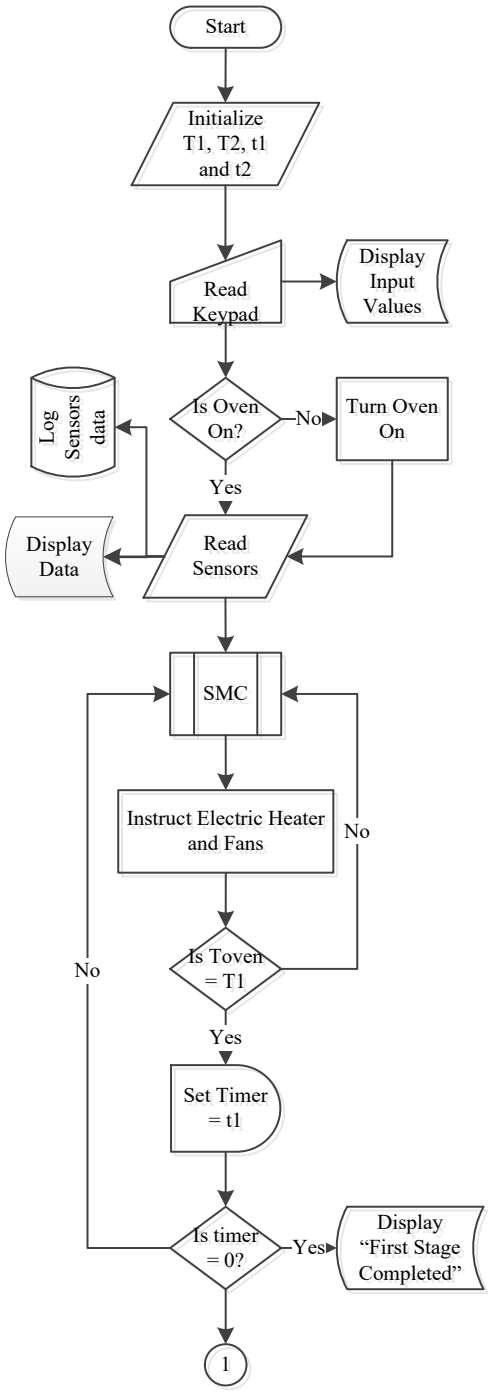


Figure 9(a): Flowchart of the controller – First stage

As the temperature reference is updated in the continued flowchart shown in Figure 9(b), the processor compares the temperature Tout, of hot extracted air from the collector to the new reference T2. If Tout is greater than T2, the electric heater is turned OFF so that the oven is solely powered by the collector; else the combined power sources continues as the SMC subroutine is called. The regulation of the electric heater, extraction and circulating fans by SMC generated signal begins to drive the oven temperature to the new reference point. Once again timer is set with the value of t2 and after the second stage soak time elapses "Paint Cured" is displayed on the LCD screen. The electric heater turned OFF and the oven shuts down.

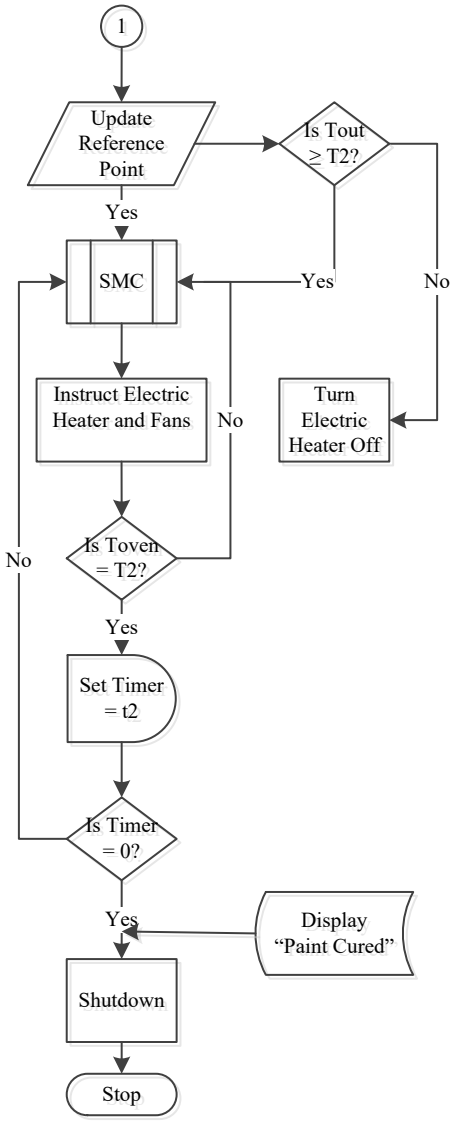


Figure 9(b): Flowchart of the controller – Second stage

Sliding mode control is a fast responds simple control method that switches between two discontinuous states in order to drive an output to track a reference input. Figure 10 is the flowchart algorithm of sliding mode control solar thermal/electricity powered oven. The control execution starts by reading three temperature variables; the reference temperature Td, this changes for a multiple reference inputs (T1 or T2 as in the main flowchart). Td is the desired temperature at any stage (T1 or T2). The second variable is Ta, which is the temperature of the surroundings and the third variable is the actual temperature in the oven, Toven.

Toven is continually acquired to track Td. As the variables are updated the equivalent control E_{s-equ} , and the switching control E_{s-sw}, are computed. The equivalent control is the control mechanism that maintains Toven to be steady once it reaches the sliding surface ($T_d - T_{oven} = 0$). It is computed by taking the quotient of the total temperature rise in the oven chamber ($T_{oven} - T_a$) and the total heat capacity of the oven $\frac{1}{R_T}$, (where $\frac{1}{R_T} = \frac{C_T}{\tau}$ and $\tau = C_T \times R_T$).

The switching control is the mechanism that drives the actual output Toven, to the reference point Td in order to reach the sliding surface. The operation is achieved by changing the state of the Toven to be opposite in direction to any deviation from the sliding surface. That is taking the product of signum function of the error signal ($sign(T_d - T_{Oven})$, the heat capacity of the oven system C_T and a positive constant M, so chosen as to be greater than the maximum limit of the system disturbances.

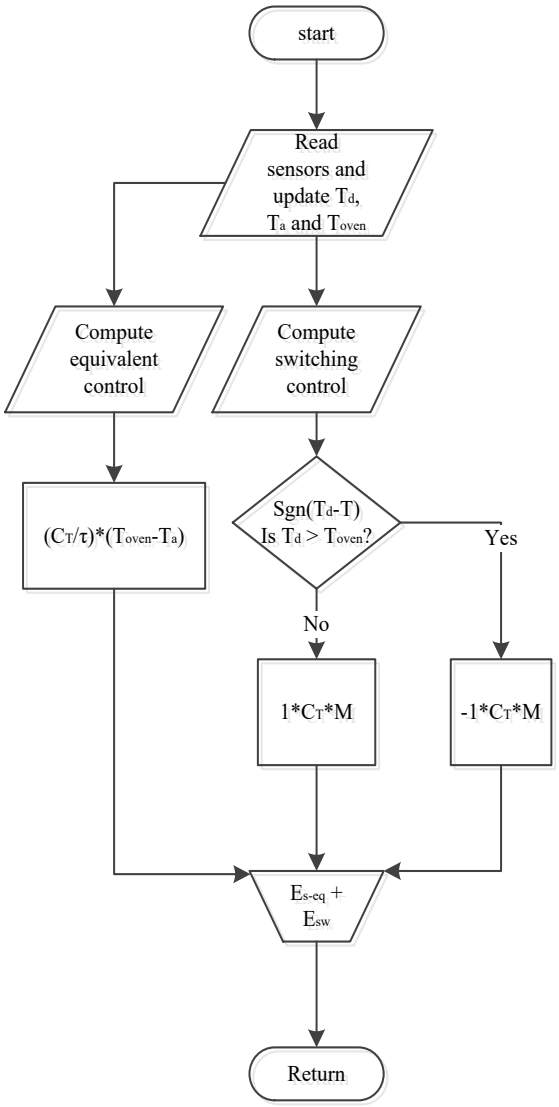


Figure 10: Sliding mode control flowchart

Table 2: Experimental test results

Quantity	Value
Rise Time	150.2369 (s)
Settling time	165.7523(s)
Settling Max.	60.8341 (°C)
Overshoot	1.7970 (%)
Undershoot	0 (%)
Peak	61.1800 (°C)
Peak Time	213.1673 (s)
Air Channel Temperature Max.	50.10 (˚C)
Hydraulic Space Temperature Max.	69.7 (˚C)
Collector Total Energy Supply	419,250.7 (W)

Signum function outputs $+1$ when the error signal is positive and -1 when the error signal is negative. The control signals from the equivalent control and the switching control are summed up in a Flowchart merger to get the control law of the sliding mode control. The control law generates the signal which is returned to the micro-controller and the process repeats until the paint is cured.

3.4. Experimental Test Results and Analysis

The system setup is for combined solar and electricity as sources of heat energy. The experiments are performed for single temperature reference and double temperature references. A plot obtained from the single temperature level test is shown in Figure 11. The plot shows the temperature evolution in thick blue line which is how the sliding mode controller drives the oven temperature to the reference temperature shown in black line.

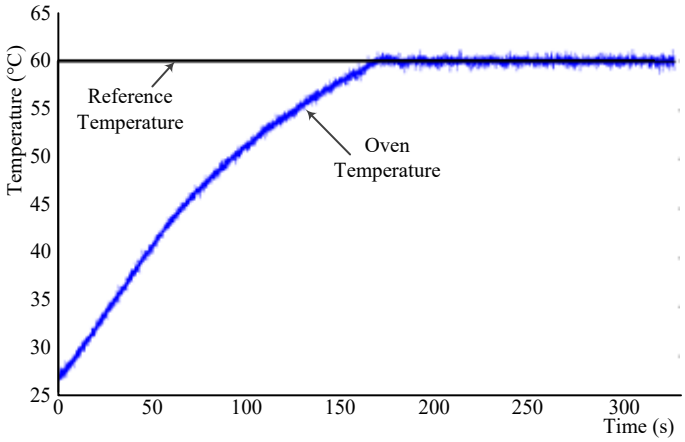


Figure 11: Oven temperature-time plot for a single reference point

Just as in the simulation results, the system has a rapid rise time in the experimental results; which demonstrates that the controller is an improvement over other convectional control methods. The controller also proves to be effective in tracking the reference temperature of 60°C with little or no deviation.

Table 2 is the statistical data collected from the analysis of oven temperature evolution logged in external drive while performing the experiment.

The table shows a rise time of 150.2369 s and settling time of 165.7523 s. It also shows settling maximum of about 60.8°C and an overshoot of about 1.8% with 0% undershoot. The maximum temperature recorded in the oven is 61.18°C which is the peak that occurred at 213.17 s.

The maximum temperature in the air channel that connects the solar collector to the oven is 50.10°C and the maximum temperature in the hydraulic space (gap between absorber plate and glass cover) of the collector is 69.7°C. The total energy supplied by the collector is 419,250.7 W.

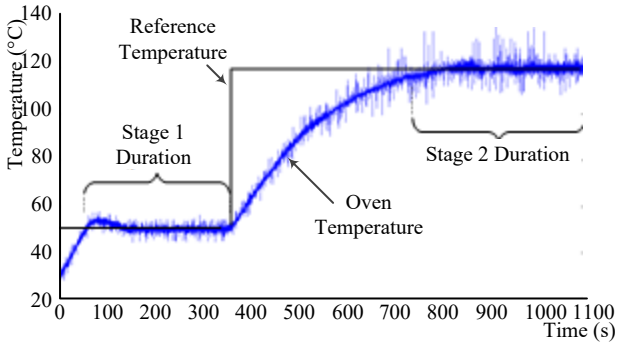


Figure 12: Oven temperature for double reference points

The second experimental demonstration test results are shown in the plot of Figure 12. The two reference points are 50°C and 120°C with each set to soak time of five minutes (300 s). The figure again shows that the controller effectively tracks multiple references at the various set time without much overshoot. However, noisy spikes begin to appear as the oven is in continuous use which become more pronounced as the temperature elevates. This is a design issue as a result of using thermistor with 150°C limit of maximum accurate range.

3.5. Payback Period and Energy Savings

Payback period is an estimate of the amount of time that is needed for cash inflow realized from a project to offset the initial cash outflow; it ignore the cost of maintenance [14-18] [19, 20, 21]. Powering of paint curing oven system with flat-plate air solar thermal collector to come with addition cost and the time required for the collector to generate energy that worth the cost.

In Table 2, the total energy the supplied to the oven by the collector is 419,259.7 W, for a period of 325 s. This is equivalent to1,290 W/s

This implies that in one hour the collector can supply

$$1290 \times 60 \times 60 = 4{,}644\ MJ \qquad (27)$$

Equation (27) shows that the collector generates 4.644 MJ of energy per hour. A kilowatt-hour of energy is 3.6 MJ and costs ₦47.8 in Nigeria that is equivalent of $0.13 (or thirteen cent).

$$\Longrightarrow ₦47.80 \times \frac{4.644}{3.6} = ₦61.66 \quad (28)$$

Equation (28) shows that the collector saves ₦61.66/hr

Assuming seven hours (7 hrs) of insolation per day over a period of one year (365 days) the equivalent cost of energy savings is ₦157,546.41 as shown in (29).

$$₦61.66 \times 7 \times 365 = ₦157{,}546.41 \quad (29)$$

The cost of solar thermal collector in Figure 8(a) and its operation auxiliaries is ₦210,009.36; the payback period is thus:

$$Payback\ period\ = \frac{total\ collector\ cost}{amount\ generated\ per\ year} \quad (30)$$

$$payback\ period = \frac{210{,}009.36}{157{,}546.41} = 1.33\ yrs \quad (31)$$

The payback period is about sixteen months.

4. Conclusion

This paper is an extension of work originally presented in 2019 IEEE International Conference on Sustainable Energy Technologies and Systems (ICSETS) [1]. It extends the control methodology approach and shows how the stability verification on the system model is done.

The work uses sliding mode control techniques and integration of solar thermal collector to solve oven temperature control and power optimization problem. Comparison of a model of an oven system designed and simulated in MATLAB/SIMULINK shows that sliding mode control method is an improvement over other control methods like PID control.

A convective paint curing oven has been built to verify the control method on paint curing process; Flat-plate air solar collector has also been built and incorporated into the oven system to reduce cost on utility bill. The developed control method was implemented for multi-stage temperature level using micro-controller and thyristor modules that switches electricity if the solar thermal collector is not supplying up to required temperature.

The system was setup and tested for verification and the experimental results show behavior similar to the simulation results. The results also show that the collector energy generation can payback the running and installation cost within sixteen months on an average insolation of 200W/m2 in Nsukka. The collector provides 43.27% of the energy requirement by the oven when the oven operates at 120˚C, which is considerably good.

Acknowledgment

This work is sponsored by Mirai Denchi Nigeria Limited

References

[1] O. V. Ajah and E. C. Ejiogu, "Solar Therma/Electricity Paint curing Oven," in *IEEE International Conference on Sustainable Energy Technologies and Systems (ICSETS)*, Bhubaneswar, India, Mar. 2019.

[2] R. Talbert, "Curing," in *Paint Technology Handbook*, Boca Raton, USA, Press Taylor & Francis, 2008, pp. 161-167.

[3] Despatch, "Industrial Oven Selection Guide," in *Thermal Processing Technology*, 8860 207th Street West Minneapolis, MN 55044 USA, Despatch, 2016.

[4] N. Hambali, A. A. Rahim and A. A. Ishak, "Various PID Controller Tuning for Air Temperature Oven System," in *IEEE International Conference Proc. on Smart Instrumentation, Measurement and Applications (ICSIMA)* , Kuala Lumpur, 2014.

[5] Y. Z. X. Gao and W. Guo, "Simulation and Research of Fuzzy Immune Adaptive PID Control in Coke Oven Temperature Control System," in *IEEE Proceedings of the 6th World Congress on Intelligent Control and Automation*, Dalian, China, 2006.

[6] K. Ohishi, "Robust Temperature Control of Thermostatic Oven Based on Adaptive and Fuzzy Algorithm," in *IEEE Industrial Electronics Society IECON 15th Annual Conference*, 1989.

[7] Z. Xuelei, W. Songling, C. Haiping and Z. Lanxin, "System Design and Economic Evaluation of Coke Oven Gas Utilization Projects," in *IEEE Conference*, 2010.

[8] F. Pask, P. Lake, A. Yang and e. al., "Pask, F., Industrial oven improvement for energy reduction and enhanced process performance," in *Clean Techn Environ Policy*, 2017.

[9] J. Kent, "http://www.pcimag.com," Apr., 2008. [Online]. Available: https://www.pcimag.com/articles/95662-a-uniform-cure. [Accessed 10 Feb., 2017].

[10] E. C. Ejiogu, "High Performance Vector Control of the Induction Motor by Application of the Variable Structure System Theory," PhD Thesis, Shinshu University, Nagano-City, Japan, 1994.

[11] V. I. Utkin, "Sliding Mode Control Design Principles and Applications to Electric Drives," *IEEE Trans. Industrial Electronics,* vol. 40, no. 1, p. 23 – 36, Feb, 1993.

[12] K. D. Young, V. I. Utkin and Ü. Özgüner, "A Control Engineer's Guide to Sliding Mode Control," *IEEE Trans. Control Systems Technology,,* vol. 7, no. 3, pp. 328-342, 1999.

[13] B. Mohamed, M. Benyounes and Z. Souhila, "Comparison of Sliding Mode Control and Fuzzy Logic Applied to Wind Turbine Emulator," in *IEEE Conference 3rd International Symposium on Environmental Friendly Energie and Applications (EFEA)*, 2014.

[14] M. Jzernik and K. Dal, "Experimental Comparison of Discrete Time Sliding Mode and Conventional, PI Current Controller for IM Drives," in *IEEE 12th International Power Electronics and Motion Control Conference*, 2006.

[15] K. Eriksson, D. Estep and C. Johnson, Applied Mathematics: Body and Soul Volume 1: Derivatives and Geometry in IR3, Singapore: Springer, 2004.

[16] V. Ajah and G. A. a. E. Ejiogu, "Sliding Mode Temperature Control for Hybrid Solar/Electricity Oven," *International Journal of Mechatronics. Electrical and Computer Technology (IJMEC),* vol. 10, no. 37, pp. 4654 - 4664, July, 2020.

[17] A. Victor, "Sliding Mode Temperature Control: Hybrid Solar Thermal/Electricity Paint Curing Oven," M.Eng Thesis, Electrical Engineering University of Nigeria, Nsukka, 2018.

[18] N. Derbel, G. Jawhar and Z. Quanmin, Applications of Sliding Mode Control, Singapore: Springer, 2017.

[19] S. Stelling, T. Y. R. Syah, R. Indrawati and D. Dewanto, "Role of Payback Period, ROI, and NPV for Investment in Clinical Health Business," *International Advanced Research Journal in Science, Engineering and Technology (IARJSET),* vol. 5, no. 7, pp. 78 - 82, July, 2018.

[20] A. F. AWOMEWE and O. Oludele Olawale, "THE IMPORTANCE OF THE PAYBACK METHOD IN," Blekinge Institute of Technology, Blekinge, 2018.

[21] M. K. Al-Ani, "A Strategic Framework to Use Payback Period in Evaluating the Capital Budgeting in Energy and Oil and Gas Sectors in Oman," *International Journal of Economics and Financial Issues (IJEFI),* vol. 5, no. 2, pp. 469 - 475, 2015.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 191-200 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Biometric System Vulnerabilities: A Typology of Metadata

Abdou-Aziz Sobabe*, Tahirou Djara, Antoine Vianou

Laboratoire d'Electronique de Télécommunication et d'Informatique Appliquée (LETIA/EPAC) Université d'Abomey-Calavi, Benin, Institut d'Innovation Technologique (IITECH) Abomey-Calavi, Bénin

ARTICLE INFO

Article history:
Received: 19 September, 2019
Accepted: 13 January, 2020
Online: 22 January, 2020

Keywords:
Biometric vulnerabilities
Metadata
Area Under the Curve (AUC)

ABSTRACT

This study presents a root cause analysis of biometric vulnerabilities and provides a comprehensive typology of metadata in biometric adaptation. Although they are more reliable and secure than traditional authentication methods, biometric techniques are subject to vulnerabilities that pose challenges. Faced with the proliferation of cases of identity theft and fraud, biometrics is increasingly used to protect assets and people in several areas such as commercial, forensic and government applications. As a first step, a metadata analysis was performed. A focus has then been placed on their role in the fight against biometric vulnerabilities. Thus, the vulnerabilities studied have been classified into two main categories: intrinsic limits and adverse attacks. Finally, one of the scenarios considered was implemented, particularly the case of the combination of skin color with facial recognition. The implementation resulted in encouraging results with an Area Under the Curve (AUC) of 0.826 for the face system and 0.908 for the multimodal system.

1. Introduction

The fast-growing digital transformation is a process of fully integrating digital technologies into all of an organization's activities. In this digital world, the authentication stage is often considered as the weak link in computer security, because of the large number of identity theft cases. User authentication is an important challenge in securing assets. In reference to computer security, biometrics refers to the use of morphological, behavioral or biochemical characteristics to determine or verify the identity of a user.

Security is about "protecting" "assets" against all forms of "threats" [1]. Threat is the type of action that may be harmful in the absolute, while vulnerability (sometimes called a gap or breach) represents the level of exposure to the threat in a particular context. Finally, the countermeasure is the set of actions implemented to prevent the threat. The countermeasures to be implemented are not only technical solutions but also training and awareness measures for users, as well as a set of clearly defined rules. From the notions of threat, vulnerability and countermeasures, we can define the notion of risk as a danger that can be sustained at any time. The risk in terms of security is generally characterized by the following equation:

$$Risk = \frac{Threat * Vulnerabilities}{Countermeasures}$$

The security procedure can be represented by Figure 1 below.

As we can see, vulnerabilities are at the heart of the security procedure. Biometric authentication systems are positioned as the means of countermeasure put in place to ensure the protection of assets. But these systems face a number of challenges and vulnerabilities. Among the solutions to these challenges and vulnerability, the use of metadata figures prominently. In the following sections, we present an overview of biometrics metadata before addressing their role in addressing identified vulnerabilities.

2. Overview of some Non-Current Biometric Features

Biometric features are more reliable and secure for person recognition than traditional methods based on knowledge or possession. They can be morphological (palmprint, palm vein, (DNA, odor) or bioelectrical (electrocardiogram, electromyogram) [2–5].

2.1. Morphological Features

2.1.1. Hand and Finger Geometry

The human hand has some relatively stable features (e.g., length of fingers) which are peculiar (although not very distinctive) to an individual. They can therefore be used for biometric recognition purposes. The "measurement" of the hand is made up of several measures such as finger dimensions, joint

*Corresponding Author: Abdou-Aziz Sobabe, Email: azizsobabe@yahoo.fr

characteristics, palm and hand shape. Such an identification system studies on average 90 lines of the hand to recognize an individual. The first step of authentication is that of the scan. To do so, the person must put his hand on a platinum. The fingers must be properly placed. etc.), behavioral (voice, signature, gait, keystroke), biochemical. An infrared camera then takes an image from two different angles to obtain a three-dimensional reproduction of the hand. It normally takes no more than three seconds to read. The hand geometry systems have large physical size, so they cannot be easily embedded in existing security systems [6,7]. The hand geometry technology is used for applications such as access control, where one is cocerned about an individual trying to gain access by using someone else's card or PIN. In terms of advantages, the hand geometry authentication technique seems less intrusive than fingerprint registration. In addition, it requires the retention of smaller files compared to a database for comparing fingerprints. For example, scanning an image of a hand is only a space of about 15 bytes while the image of the fingerprint represents a file size of around 500 bytes or more. Finally, external factors such as the moisture of the skin and dirt on the hand do not prevent a good measurement. The same is true for burns and minor cuts. Compared to the fingerprint, this technique has a high rate of false positives. This means that family members with significant physical similarities can easily deceive the system. In addition, the shape of the hands is likely to change due to diseases related to old age, such as arthritis. Furthermore, the geometry of the hand requires a larger scanner than a fingerprint reader. As a result, its use becomes embarrassing when it comes to securing a small object such as a computer [8].

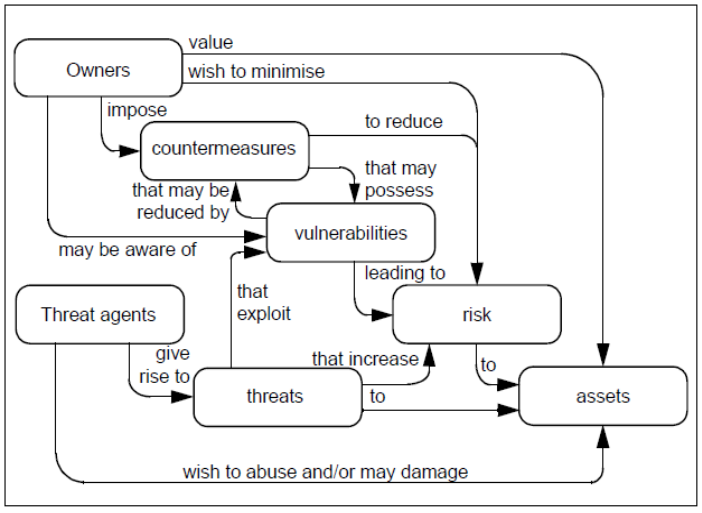


Figure 1: Security procedure [1]

2.1.2. Iris

Iris scan technology is one of the most reliable biometric recognition systems developed recently. The recognition of the iris is a biometric technique to recognize an individual by the observation of his iris. The iris is located in the aqueous humor, it is surrounded by the white of the eye, the pupil is located in its center, the cornea is in front of him and the crystalline behind. The iris corresponds to the colored part of the eye and it is this part that is used in biometrics. Indeed, the texture of the iris (i.e. the pattern of the iris), includes many features. Those most often used in biometrics are stripes, pits, and furrows. These elements of the iris remain fixed, they vary only very little during a lifetime. Each pattern determined by the chaotic processes during embryonic development is stable and unique (the probability of similarity is 1 in 10 power 72) [8,7]. In addition, the pattern of the iris is not genetic in contrast to the color of the eyes. So, two individuals, even if they are parents, can have the same iris color but never the same motive. That's why the iris makes it possible to distinguish the identical twins. Iris-based systems have a very low False Accept Rate (FAR) compared to other biometric traits. Nevertheless, it is important to note that iris systems have a high False Reject Rate (FRR) [6,9]. Iiris sample images quality have an impact on the accuracy of Visible Wavelength iris recognition systems. To deal with this problem, several iris image quality enhancement methods have been developed by researchers. Liu and Charrier [10] evaluated the performance of nine image enhancement approaches on Visible Wavelength iris biometric images. Experimental results showed that three of them (CLAHE, MEDIAN, and UNSHARP) can highly improve the system performance. In contrast, two other methods (SSR and WIENER) have a lower level of performance improvement.

2.1.3. Retina

The retina is the sensory membrane that lines the inner surface of the back of the eyeball. It's composed of several layers, including one that contains specialized cells called photoreceptors (https://www.allaboutvision.com/resources/retina.htm). The elements that distinguish two retinas are the veins that line them. The arrangement of these veins is stable and unique from one individual to another (from one eye to another). And the models that come from inherit the stability of this disposition. The difficulties involved in capturing the image of a retina are as much psychological as medical and technical. To obtain an image of a retina, it is necessary to illuminate the back of the eyeball by means of a light beam. This beam is of very low intensity so as not to disturb the user; it is safe and of lower intensity than on ophthalmic devices. A very precise camera system then comes to recover the image of the retina. Retina readers are available, and provide a very high level of security. After capturing an image of the retina, the reading device software cuts a ring around the fovea. In this ring, it locates the veins and their orientation. It then creates an "eye signature" that is used for retinal recognition. The operation itself is quite simple to describe but the algorithms remain relatively complex.

Retinal biometry provides a high level of recognition. This technology is well suited for high security applications (military sites, safes, etc.). The arrangement of the veins of the retina ensures a good reliability and a high barrier against fraud. [11].

2.1.4. Palmprint

Several authors have researched palm-based biometric recognition. These include [12,11,13]. Palm print recognition is a biometric authentication method based on the unique patterns of various characteristics in the palms of people's hands. It is a relatively new biometric when compared to other biometric systems such as face, fingerprint and iris recognition systems. Palmprint is concerned with the inner surface of a hand. A palm is covered with the same kind of skin as the fingertips and it is larger

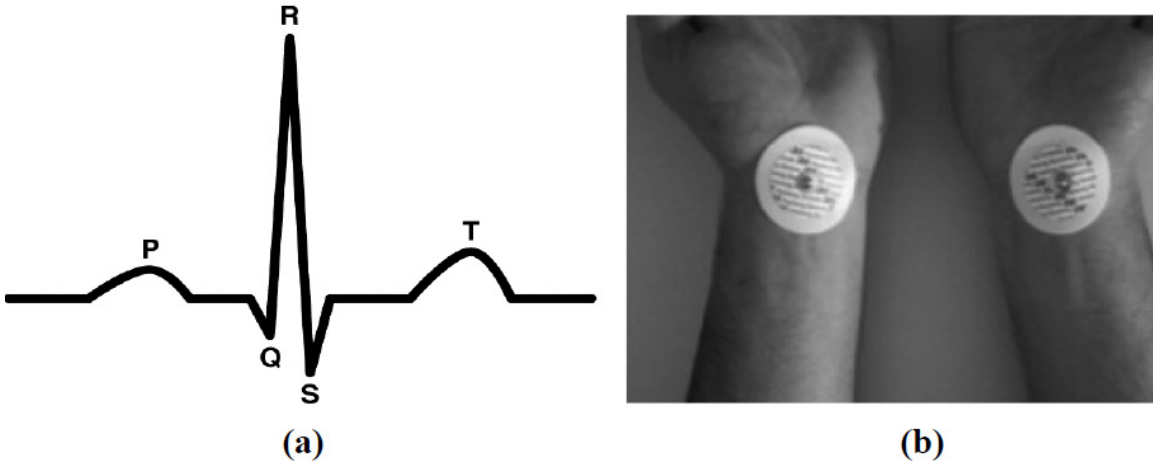


Figure 2: ECG biometrics: (a) ECG signal with regular rhythm (b) positioning of the electrodes on the forearms for ECG capture [14]

than a fingertip in size. Five features of a palmprint are usually used to uniquely identify a person. It is: (a) Geometry Features (such as width, length and area); (b) Principal Line Features (both location and form of principal lines); (c) Wrinkle Features (the thinner and more irregular lines); (d) Delta Point Features (the center of a delta-like region in the palmprint, usually located in the finger-root region) and (e) Minutiae Features (see fingerprint).

Other morphological traits are often used in biometric recognition. Those systems include Palm Vein, Ear and Facial Thermogram. More information can be found in [12,4].

2.2. Behavioral Features

2.2.1. Gait

Gait refers to the way a person walks, and is one of the few biometric traits that can be used to recognize people at a distance. Therefore, this trait is very appropriate in surveillance scenarios where the identity of an individual can be surreptitiously established [11]. Gait-based systems also offer the possibility of tracking an individual over an extended period of time. However, the gait of an individual is affected by several factors including the choice of footwear, nature of clothing, affliction of the legs, walking surface, etc. [9]. This type of biometric recognition system identifies individuals by their posture and the way they walk through unique characteristics that are perceivable at a low resolution, hard to conceal and non-invasive. [8].

2.2.2. Keystroke

Keystroke is a biometric solution "Software Only". It is applied to the password, which becomes much more difficult to "imitate". During the implementation of this technique, the user is asked to enter his password a dozen times in a row. Using an algorithm that exploits the support time on each key and the time between each key, the ten input is "averaged" to build a "typing profile" of the user that will serve as a reference. At the following accesses, the password entry will be coupled to a keying profile that will be compared to the reference profile. The access right is then granted according to the level of similarity of this profile with the reference. Depending on the degree of filtering that an administrator will have defined, this access will be more or less difficult. Whether additional security to an application, a Single Sign On, Intranet or Internet, in addition to security to a personal smart card, it is a solution, reliable, easy to implement and very competitive because without hardware. Its main advantages are related to rapid implementation for many users. It also significantly reduces the need for password changes and solicitation of IT services. As a major drawback, you must be in good condition before using the system at the risk of having your own password refused [8,11].

2.3. Bioelectrical Features or Hidden Features

2.3.1. Electrocardiogram

An electrocardiogram (ECG) is a test that studies how the heart works by measuring its electrical activity. At each heartbeat, an electrical pulse (or "wave") passes through the heart. This wave causes the heart muscle to contract so that it expels the blood from the heart [5]. The ECG is mainly used in clinical applications to diagnose cardiovascular disease. The ECG signal is characterized by the shape of its beats consisting of five typical waves, namely P, Q, R, S, and T or sometimes the U wave (Figure 2) [14]. ECG biometrics has been the subject of a number of studies [15–17]. The use of ECG in biometrics is relatively new. In fact, there are several biometric methods based on the ECG. There are approaches that are based on ECG analysis [18]. Others rely on the integration of analytic and appearance features extracted from ECG signals [19]. Several research studies confirm that the cardiac profile is specific to each individual and does not vary with age or even if the rhythm of the heart is racing after an effort or an emotion. Compared with other biometric modalities, the ECG is more universal and difficult to forge [20].

2.3.2. Electromyogram

Electromyogram (EMG) signals are bioelectrical signals recorded in the muscles. They provide various information on the state of the peripheral nerves. The EMG signal has several clinical applications. Its use as a hidden biometric modality can be particularly interesting. In this context, some recent experiments have been carried out [21,22,19]. In particular, this work has focused on the analysis of surface electromyography (SEMG) signals [23]. When acquiring these signals, individuals are encouraged to apply manual pressure of a constant intensity to a force probe for several seconds (Figure 3) [14]. The signal thus

obtained is analyzed in the spectral domain. Then, parameters are extracted such as the signal strength, the average frequency, the flattening coefficient and the asymmetry coefficient. Indeed, these parameters provide a device vector that we can use to characterize individuals.

In addition, it should be noted that magnetic resonance imaging (MRI) and X-rays are part of the hidden biometric techniques used for the authentication and identification of individuals [24].

3. Applications of Biometrics

The requirements related to the level of development of humanity and the security constraints require a rapid and reliable user authentication. Biometrics is an emerging field where technology improves man ability to identify a person [25]. The applications of biometrics can be divided into three main groups [26,13,27]:

- Commercial applications involve computer network access, electronic data security, e-commerce, Internet access, credit card, physical access control, cell phone, medical records management, remote study, etc One of the most popular technologies for ECG-based biometrics is the HeartID developed by researcher Foteini Agrafioti of the Canadian University of Toronto. This technology can be integrated into any electronic device, starting with smartphones, tablets or game consoles. The idea is to offer both an identification system to secure this type of device, but also to share them by activating user profiles opening on specific settings and contents. The sensors can be implanted so that the user only has to hold the terminal so that his heart rate is analyzed by the recognition algorithm, the identification taking about 1.2 seconds. A preliminary learning phase is necessary to record the electrocardiogram whose profile can then be stored in the terminal itself or on a database to which it will connect. Only requirement, the user must hold the terminal with both hands so that the signal can be captured because the electrocardiogram results from an electrical activity requiring two reference points on either side of the heart.

Another application of ECG-based biometrics is the Nymi bracelet developed by the Canadian company Bionym. This bracelet is able to detect the electrical signals emitted by the heart to control the identity of the wearer. Once identified, the owner can open his car, unlock his laptop or phone, or make payments such as with a credit card. With each electrical impulse, the heart will contract, and produce regular beats (between 60 and 80 per minute at rest). It is by analyzing this "cardiac signature" that the Nymi bracelet can identify its owner then, through a Bluetooth connection, interact with electronic devices to enter a PIN or unlock a user session.

- Government applications include national identity card, driver's license, social security, border control, passport control, voter registration, etc. For example, the Government of Benin has created a Permanent Electronic Electoral List, the first draft of which was put online in January 2015. Since then, it has served as a basis for the various elections held in the country. The Schiphol Privium device at Amsterdam Airport uses an iris sensor to speed up the process of checking passports and visas.

- Forensic applications include body identification, criminal investigation, terrorist identification, parenthood determination, missing children etc. The Federal Bureau of Investigation (FBI) of the US Department of Justice utilizes the Integrated Automated Fingerprint Identification System (IAFIS), a 10-fingerprint matching system that captures rolled prints. In 2008, the Chinese Police adopted an Integrated Automated Biometric Identification System (IABIS) solution to allow forensic fingerprint examiners the ability to cross check inmate identities for possible matches within the database [28].

4. Typology of Metadata

Metadata or ancillary information is non-biometric data that is either combined with pure biometric data (fingerprint, face, iris, signature, etc.) or exploited by a system that makes biometric adaptation. The goal of adaptation is to provide ways to deal with the challenges of biometric systems.

In general, there are two main methods for adapting biometric systems, the first link to the user (soft biometrics) and the second link to the acquisition system. In the first case, there are a multitude of traits from the user that can be extracted and combined with pure biometric modalities. Jain et al. were among the first to explore this area. They did experiments on the following traits: Height, gender, ethnicity. Later, Dantcheva et al. [29] conducted a study on several soft biometric traits such as: age, skin color, hair color, eye color, marks, etc.

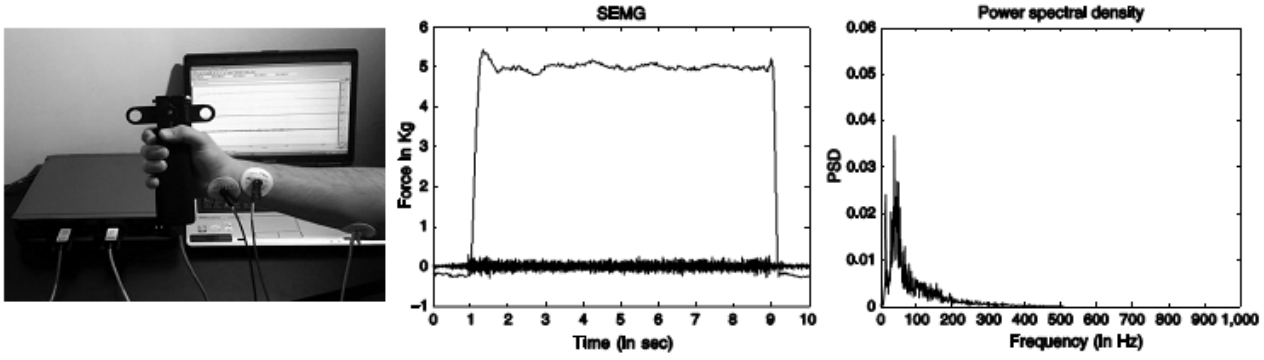


Figure 3: EMG biometrics: (a) Acquisition of an EMG signal (b) The intensity applied by the user and the relative EMG (c) EMG periodogram [14]

It should be noted that depending on the authors, certain traits are classified either as pure biometric modalities or as soft biometrics. This is the case of the gait which is sometimes considered as a behavioral modality [8] and sometimes as a trait of soft biometrics [29].

At the level of the acquisition system, adaptation can be done on the one hand at the sensor (volume, flash, etc.) and on the other hand taking into account the context of implementation (brightness, noise, etc.). Here, the metadata are not directly merged with the pure biometric data. The consideration is made at two levels. Either the system administrator makes adjustments to the various parameters such as volume, ambient brightness, flash, etc. On the other hand, the acquisition device is pre-sized to ensure automatic control of these different parameters. Such a device will be smart enough not to trigger for example a facial image when the brightness is not at a minimum required.

Several studies have focused on the role of metadata in improving biometric performance. In [30], the authors show the benefits of using gender, ethnicity and height information of the user in addition to fingerprint. The use of these soft biometric data leads to an improvement of approximately 5% over the main biometric system. [29] spoke of similar results which reduced the total error rate to 1.5% from 3.9% when body weight score is fused with fingerprint score. The table 1 below presents a typology of metadata in the biometric adaptation.

In the next section, special emphasis will be placed on the effect of metadata in the fight against biometric vulnerabilities.

5. Role of Metadata in Facing Biometric Challenges and Vulnerabilities

The challenges of biometric systems are essentially at the level of limits in terms of performance and acceptability. In addition, biometric authentication faces vulnerabilities that have been identified at eight levels on a generic biometric architecture [31]. Figure 4 shows the eight possible locations of attacks in a generic biometric system. [32] then listed a total of four different levels of biometric system vulnerabilities. These are intrinsic failure, administration issues, non-secure infrastructure, and biometric overtness. The last three levels have been grouped into adversary attacks from outside the system. These four levels of vulnerability were synthesized through a fish-bone model (see Figure 5).

5.1. Intrinsic Failure

An intrinsic failure is a biometric system security failure that generates an incorrect decision taken in a native way ie without without the intervention of an external element. It can occur even in the absence of an explicit effort by an attacker to bypass the system. This type of failure is also known as "zero-effort-attack". A biometric verification system can make two types of errors in decision making. These are false acceptances and false rejections. In the first case, false acceptances are usually caused by a lack of uniqueness in the biometric trait, which can lead to a great similarity between the characteristics of different users. At this level, adding metadata can be effective in combating this problem. Among the most probable cases, there are those of two identical twins or brothers who have a perfect similarity of the face for example. Taking into account soft biometric data such as marks and make-up will make it possible to distinguish between two individuals even if they have exactly the same face because the face recognition algorithms do not take into account these accessory data.

A legitimate user may be falsely rejected by the biometric system because of the large differences between the user's stored model and the input biometric feature sets. These intra-user variations may be due to an incorrect interaction of the user with the biometric system. This is the case, for example, of the effect of twists on a fingerprint image or the effect of the changes of pauses on a face image. Metadata do not provide a specific solution for these cases of intra-user variations. But increasingly, the biometric verification algorithms take into account these issues. [33] have proposed a contactless fingerprint recognition algorithm that addresses this issue. On another level, it should be noted that the false rejections can be due to noise introduced at the sensor such as, for example, the residual traces left on a fingerprint sensor. Metadata does not provide a solution to this limit. Multimodal

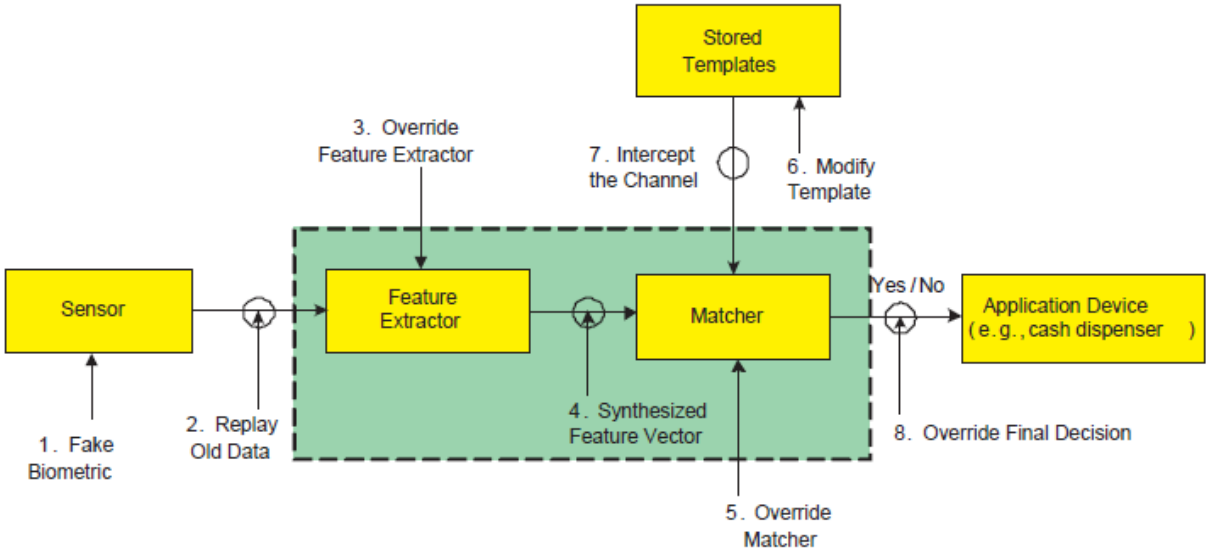


Figure 4: Possible locations of attacks in a generic biometric system [32]

biometrics offers an alternative to false rejects. Indeed, when one of the modalities is affected of noise, the system is able to ignore it and pass to the following modality.

For the limits related to the detection technology, the exploited metadata can be the volume (case of the voice modality) or the flash (case of the face, the contactless fingerprint, etc.). Finally talking about limits in the context of implementation, parameters such as brightness, noise, and temperature are used as metadata to deal with intrinsic errors.

5.2. Adversary Attacks

An adversary intentionally launch an attack on the biometric system by exploiting a multitude of loopholes in the system design. The three possible forms of attack are: Administration attacks, non-secure infrastructure attacks, and biometric overtness attacks.

5.2.1. Administration Attack

This attack is still called insider attack. It usually comes from malicious, familiar people who exploit vulnerabilities created by poor biometric system administration. These include (i) the integrity of the registration process (for example, the validity of credentials submitted during registration), (ii) collusion (or coercion) between the opponent and the system administrator or a legitimate user and (iii) abuse of exception handling procedures. Under these conditions, the use of metadata will not provide a response to this type of attack in the absolute. On the other hand, the additional use of soft biometric traits can make the system more complex and require more effort on the part of the attacker.

5.2.2. Non-Secure Infrastructure

The infrastructure of a biometric system includes the hardware, software and communication channels between the different modules. An opponent can manipulate the biometric infrastructure in several ways, which can lead to eight security breaches as described below.

First level:

Falsified biometric data: A reproduction of the biometric data used will be presented to the biometric sensor. In the case of authentication by the fingerprint, the attacker can present a false finger facing a contact-based sensor or just have the image of a finger facing a contactless sensor. It should already be noted that this form of attack is more common with the fingerprint which is a widely used modality with a high discriminatory power. On the other hand with other modalities such as face, retina, and signature, it is more difficult to present a falsified data. Moreover, if the biometric system integrates a soft biometric feature to be combined with pure biometric modality, this form of attack becomes more complicated to implement. Take for example a case of fingerprint authentication associated with the height of the user. Even if the fingerprint is falsified, the impostor will be stucked during the enrollment of the height. Considering another case of face authentication combined with the skin color automatically extracted from the facial image, an impostor will have a hard time tampering with a face while taking into account the specificities and requirements related to the detection of the skin color.

Second level:

Transmission of intercepted biometric data: Here, the attacker plays back an old biometric data stored in the system without passing through the biometric sensor. This is the case with the presentation of an old copy of the fingerprint image. Since the attacker bypasses the biometric sensor by providing the system with an old recorded data, the metadata will have no effect against this form of attack.

Third level:

Attack on the feature extraction module: This module could be replaced by a Trojan horse so as to produce information chosen by the attacker. The legitimate user does not realize that this module has been corrupted and provided information in accordance with the hacker's instructions. The feature extraction module being compromised by the hacker, the metadata will not be effective against this kind of attack.

Fourth level:

Alteration of extracted features: After data is obtained by the feature Extraction Module, it is altered or even replaced by other data defined by the attacker. For non-secure infrastructure attacks, from the second level of attack to the eighth level, we are in situations where the biometric system is corrupted and will only provide responses according to the intent of the hacker. Metadata will not be effective in these contexts.

Fifth level:

The matching module is replaced by a malicious module: This module could be replaced by a Trojan horse to artificially produce high or low scores.

Sixth level:

Database corruption: The database of biometric templates is available locally, remotely or distributed on multiple servers. In this type of attack, the attacker modifies one or more models to allow an impostor or even to prevent a legitimate user from accessing it.

Table 1: Typology of metadata in biometric adaptation

Level of adaptation	User	Acquisition system	
Type of adaptation	Soft biometric	Sensor	Context
Example	Marks; Make-up; Age; Weight	Volume (microphone); Flash (camera)	Brightness, noise, temperature
Exploited data	Metadata		
Training need	Without training (immutable)		After training

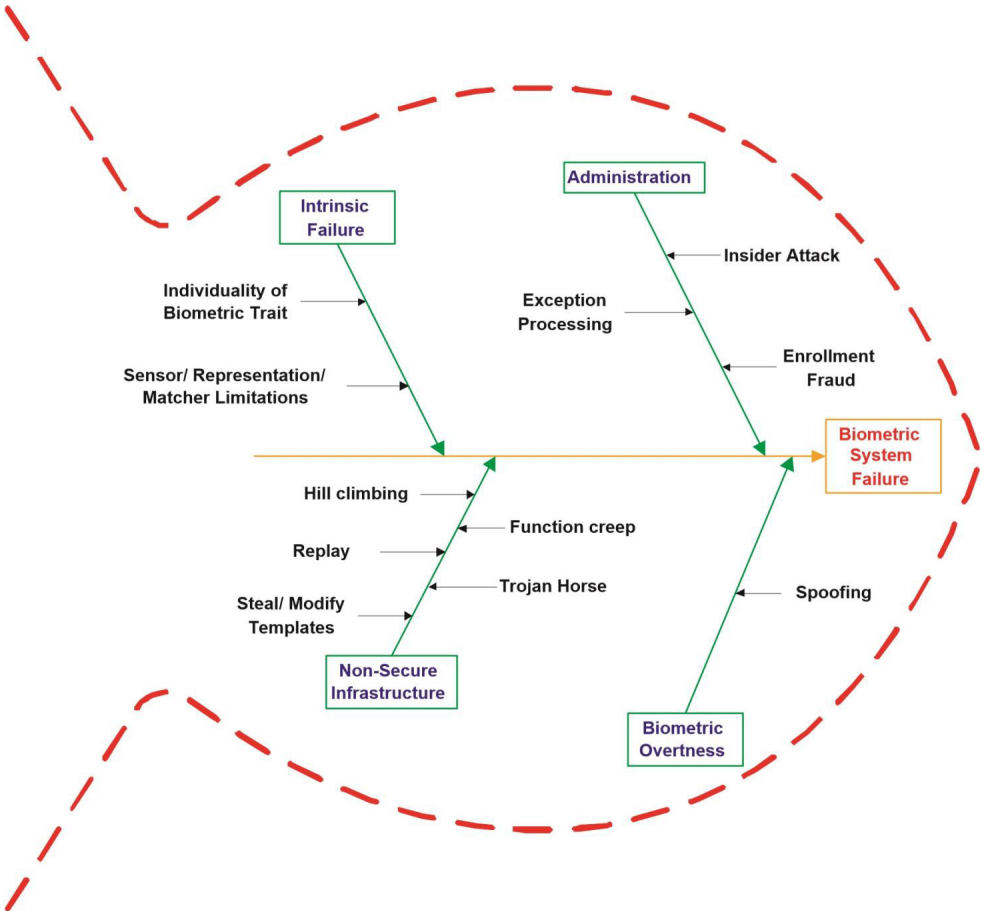


Figure 5: Fish-bone model for categorizing biometric system vulnerabilities [32]

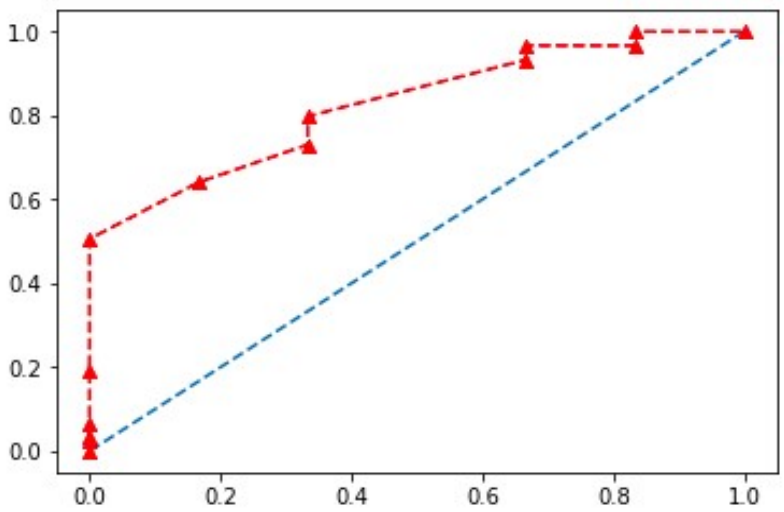


Figure 6: ROC curve with AUC: 0.826

Seventh level:

Attack on the channel between the database and the matching module: In this type of attack, the models are altered on the transmission link connecting the model base and the matching module.

Eighth level:

Alteration of decisions (accepted or rejected): This type of attack alters the boolean decision (yes or no) made by the decision module. The danger of this attack is high because even if the system is robust in terms of performance, it was rendered useless by this type of attack.

5.2.3. Biometric Overtness

An attacker may discreetly acquire the biometric characteristics of a legitimate user and use them to create an artificial production of the biometric feature (fingerprints taken from the surface of a contact-based sensor). Therefore, if the biometric system is not able to distinguish between a live biometric presentation and an artificial parody, an opponent can circumvent the system by presenting falsified traits. This case perfectly matches the first level of attack on non-secure infrastructure. The analysis performed at this level is also valid here. It should be noted that soft biometrics represents a response to this form of attack for two reasons. Indeed, the user is obliged to appear in person in front of the data acquisition device or this data is automatically extracted from a pure biometric modality having a number of characteristics hardly respectable by the artificial productions.

The table 2 below summarizes for each vulnerability category the existence of effects as well as the possible actions of the metadata. Examples of usable metadata have been listed at different levels in order to provide more probable scenarios of integration of multi-biometric metadata using a variety of disparate fusions of traits.

6. Skin Color as a Solution to Overtness in Face Authentication

6.1. Materials and Methods

The objective of this section is to present a practical case of adding metadata to a pure biometric modality in order to fight against certain vulnerabilities. The materials and methods used are among the most recent technologies in progress. The implementation of the proposed system was carried out under Python framework with the OpenCV package (https://sourceforge.net/projects/opencvlibrary/). Two stages are to be considered in accordance with the two modalities that are the face and the skin color (of the face). First, we performed face recognition using the LBPH (Local Binary Pattern Histogram) algorithm [34]. Then, for authentication by skin color, the Haar Cascade algorithm [34] made it possible to detect the face, while the extraction of the dominant skin colors was carried out automatically using the algorithm of the K-means [35]. Only one database was used; it is the Caltech faces database (http://www.vision.caltech.edu/html-files/archive.html). This database contains 450 images of faces of size 896x592 in JPEG format for 27 unique people with a variable number of images. For experimental purposes, we have resized the database to 19 people with 17 images each. 12 images were used to form the model and 5 images were used for the tests.

6.2. Experimental Results

After implementing these two modalities, the matching scores obtained were merged using the weighted sum method. A weight of 0.9 was assigned to the modality of pure biometrics (face) while the modality of soft biometrics (skin color) received a weight of 0.1. The experimental results are satisfactory with a performance of 82.6% for the facial system and 90.8% for the multimodal system. Figures 6 and 7 respectively show the results (ROC curve) of the face alone and of the face combined with the skin color.

6.3. Discussion

The inclusion of metadata in face authentication allows on the one hand to improve recognition performance and on the other hand to make the system less vulnerable. The calculation of the processing time for the two modalities gave the following results. For the face, the average processing time (for 15 images) is 1.24 s. The average time for skin color is 7.21 s. The fusion of the facial and skin color scores takes an average time of 1.52 s (always for the 15 images). It should be noted that the total processing time of the multimodal system increases due to two factors, namely the extraction time of the skin color and the time taken to merge the two scores. This confirms once again that each biometric authentication system has its advantages and limits. These are the application cases and the implementation conditions which generally determine the choices to be made.

7. Conclusion and Future Work

This paper presented a detailed typology of metadata on multi-biometric system vulnerabilities. We have provided an overview of some non-current biometric sensing systems before dealing with the applications and privacy issues. Furthermore, we discussed the categorization of biometric metadata, namely user level (ie soft

Table 2: Typology of metadata on biometric system vulnerabilities

Vulnerability category	Intrinsic failure	Adversary attacks		
		Administration Attack	Non-secure infrastructure	Biometric overtness
Existence of metadata effect	Yes (except the case of false rejects)	Partial	Yes (1 case out of 8)	Yes
Possible actions	1- Reduce false acceptances 2- Limit FTE and FTA	Make the system harder to attack	Prevent attack by presenting falsified biometric data	Prevent attack by presenting falsified biometric data
Example of metadata	1- Marks (tattoos, scars, etc.); make-up 2- Volume, flash; Brightness, noise, temperature	Weight, body shapes, etc.	Height, eye color, etc.	Gait, gender, skin color, etc.

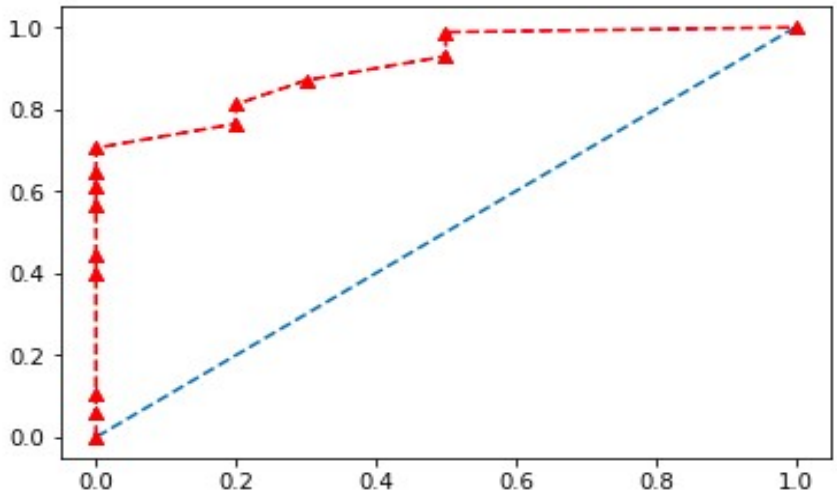


Figure 7: ROC curve with AUC: 0.908

biometric) and acquisition system level (ie sensor and acquisition context). To finish, a practical case of adding metadata to a pure biometric modality have been presented in order to fight against face recognition vulnerabilities.

Contactless fingerprint authentication [33] is a promising technique but one that is potentially vulnerable on several levels. For future work, we will explore the role of metadata in addressing contactless fingerprint vulnerabilities.

References

[1] [Common Criteria, 2012] Common Criteria for Information Technology Security Evaluation ; Part 1: Introduction and general model ; September 2012 ; Version 3.1 ; Revision 4

[2] M. Nageshkumar, P.K. Mahesh, and M.N. Shanmukha Swamy, "An efficient secure multimodal biometric fusion using palmprint and face image," International Journal of Computer Science Issues, Vol. 2, pp 49-53, 2009, arXiv:0909.2373v1.

[3] A. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," Elsevier, Pattern Recognition 38, pp 2270-2285, 2005, DOI:10.1016/j.patcog.2005.01.012.

[4] R. Gad, A. El-Sayed, N. El-Fishawy, and M. Zorkany, "Multi-biometric systems: A state of the art survey and research directions," International Journal of Advanced Computer Science and Applications, Vol. 6, No. 6, pp. 128-138, 2015, DOI : 10.14569/IJACSA.2015.060618.

[5] S. N. Abbas, M. Abo–Zahhad, S. M. Ahmed, M. Farrag, "Heart-ID: human identity recognition using heart sounds based on modifying melfrequency cepstral features," © The Institution of Engineering and Technology, IET Biom., Vol. 5, Iss. 4, pp. 284-296, 2016, doi: 10.1049/iet-bmt.2015.0033.

[6] A.K. Jain and A. Kumar, "Biometric recognition: An overview," Chapter 3, © Springer Science + Business Media B.V., pp 49-79, 2012, DOI 10.1007/978-94-007-3892-8_3.

[7] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar, Handbook of Fingerprint Recognition, © Springer-Verlag London Limited, 2009.

[8] M. O. Oloyede and G.P. Hancke, "Unimodal and multimodal biometric sensing systems: A review," IEEE Access, Volume 4, pp 7532-7555, 2016, Digital Object Identifier: 10.1109/ACCESS.2016.2614720.

[9] A. Ross, K. Nandakumar, and A.K. Jain, Handbook of Multibiometrics, Springer, New York, USA, 1st edition, 2006.

[10] X. Liu and C. Charrier, "Can image quality enhancement methods improve the performance of biometric systems for degraded visible wavelength iris images?," BioSMART 2019 Proceedings, 3rd International Conference on Bio-engineering for Smart Technologies, Paris, 24th-26th April pp. 119-122, 2019, 978-1-7281-3578-6/19/$31.00 ©2019 IEEE.

[11] H. AlMahafzah and M.Z. AlRwashdeh, "A survey of multibiometric systems," International Journal of Computer Application, volume 43 No 15 April, pp 36-43, 2012, DOI : 10.5120/6182-8612.

[12] D. Cherifi, R. Kaddari, H. Zair, and A. Nait-Ali, "Infrared face recognition using neural networks and HOG-SVM," BioSMART 2019 Proceedings, 3rd International Conference on Bio-engineering for Smart Technologies, Paris, 24th-26th April, pp. 132-136, 2019, ©2019 IEEE, 978-1-7281-3578-6/19/$31.00.

[13] A.K. Jain, P.J. Flynn, and A. Ross, Handbook of Biometrics, © Springer Science + Business Media, LLC, 2008.

[14] H. Toufik, "Reconnaissance Biométrique Multimodale basée sur la fusion en score de deux modalités biométriques : l'empreinte digitale et la signature manuscrite cursive en ligne," Ph.D Thesis, Badji Mokhtar University-Annaba, 2016.

[15] Q. Zhang, "Wavelets networks: The radial structure and an efficient initialization procedure," Technical Report of Linköping University, 1992, LITH-ISY-I-1423.

[16] J. Sjöberg , Q. Zhang, L. Ljung, A. Benveniste, B. Deylon, P.Y. Glorennec, H. Hjalmarson, and A. Juditsky, "Nonlinear black-box modeling in system identification: Unified overview," Automatica, Vol.31, no.12, pp.1691-1724, 1995.

[17] A. Juditsky, "Wavelet estimators: Adapting to unknown smoothness," Technical Report IRISA, June 1994.

[18] Q. Zhang, "Using Wavelet Network in Nonparametric Estimation," Publication Interne 833, IRISA, June 1994.

[19] Q. H. Zhang, "Using Wavelet Network in Nonparametric Estimation," IEEE Trans.Neural Networks, Vol.8, pp.227-236, 1997.

[20] S. Chantaf, "Biométrie par signaux physiologiques," Ph.D Thesis, Université Paris-Est, 2011.

[21] J. Zhang, G. G. Walter, Y. Miao, and W. N. W .Lee, "Wavelet Neural Networks for function learning" IEEE Trans. signal process, Vol.43, No.6, pp.1485-1497, June 1995.

[22] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet transform," IEEE Trans. Pattern Anal. Machine Intell., Vol.11, pp.674-693, July 1989.

[23] N. Draper and H. Smith, "Applied regression analysis, Series in Probability and Mathematical Statistics," Wiley, 1981, Second edition.

[24] A. Nait-ali, "Hidden biometrics: towards using biosignals and biomedical images for security applications," 7th international Workshop on Systems, Signal Processing and their Applications, Invited paper, Tipaza, pp. 352–356, 2011, DOI: 10.1109/WOSSPA.2011.5931509.

[25] L. Allano, La Biométrie multimodale : stratégies de fusion de scores et mesures de dépendance appliquées aux bases de personnes virtuelles, PhD thesis, Institut National des Télécommunications dans le cadre de l'école doctorale SITEVRY en co-accréditation avec l'Université d'Evry-Val d'Essonne, 2009.

[26] A.K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," IEEE Transactions on circuits and systems for video technology, vol. 14, no. 1, January, pp 4-20, 2004, DOI: 10.1109/TCSVT.2003.818349.

[27] S. Guerfi, "Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D," PhD thesis, Université d'Evry-Val d'Essonne, 2008.

[28] T. K. Ho, J. J. Hull, and S. N. Srihari, "Decision combination in multiple classifier systems," IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(1), January 1994, pp 66–75, DOI: 10.1109/34.273716.

[29] A. Dantcheva, C. Velardo, A. D'Angelo, J. L. Dugelay, Bag of Soft Biometrics for Person Identication : New trends and challenges, Multimed Tools Appl 51, © Springer Science+Business Media, LLC 2010, pp. 739–777, 2011, https://doi.org/10.1007/s11042-010-0635-7.

[30] A.K. Jain, K. Nandakumar, X. Lu, U. Park, "Integrating faces, fingerprints, and soft biometric traits for user recognition," Proceedings of Biometric Authentication Workshop, LNCS 3087, Prague, pp. 259-269, May 2004, DOI:10.1007/978-3-540-25976-3_24.

[31] M. El-Abed, C. Charrier, "Evaluation of biometric systems," chapter 7. In: New Trends and Developments in Biometrics, pp. 149–169, 2012, https://doi.org/10.5772/52084.

[32] A.K. Jain, K. Nandakumar, A. Nagar, "Biometric template security," in EURASIP Journal on Advances in Signal Processing, Special Issue on Biometrics, pp 1-20, January 2008, DOI:10.1155/2008/579416.

[33] T. Djara, M.K. Assogba, and A. Vianou, "A contactless fingerprint verification method using a minutiae matching technique," International Journal of Computer Vision and Image Processing, Volume 6, Issue 1, pp 12-27, January-June 2016, DOI: 10.4018/IJCVIP.2016010102.

[34] L. Dinalankara, "Face detection and face recognition using open computer vision classifiers", Robotic Visual Perception and Autonomy, Faculty of Science and Engineering, Plymouth University, august 4, 2017.

[35] C. C. Aggarwal, C. K. Reddy, Data clustering : algorithms and applications, Taylor & Francis Group LLC, 2014.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 211-214 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

System of Sequential D-Optimal Identification for Dynamic Objects in Real-Time Scale

Fatuev Victor*, Mishin Anton

Department of Computer Technology, Tula state university, 300012, Russia

ARTICLE INFO

Article history:
Received: 14 January, 2020
Accepted: 06 March, 2020
Online: 13 March, 2020

Keywords:
D-optimal identification
real time scale
estimation of unknown parameters
technical complex
synthesis of the testing signal

ABSTRACT

The article is devoted to the formalization and development of the theory of sequential experimental design for optimal identification of dynamic objects. The algorithmic and technical support for a system of sequential D-optimal identification for a wide class of dynamic objects in real time is considered. This system is universal and can be used for optimal identification of real objects and continuous technological processes.

1. Introduction

This paper is an extension of work originally presented in 8th Mediterranean Conference on Embedded Computing (MECO) [1].

The construction of adequate models of the dynamic objects under study is an important step in solving any applied task. With a known or given structure of the dynamics model, the problem of experimental statistical parametric identification is usually solved. The quality of the resulting dynamic regression model (DRM) is determined by the probabilistic properties of the estimates for the unknown model parameters and its predictive properties.

If active influence at the input of the object is allowed, then the optimal parametric identification task can be formulated and solved, the purpose of which is to obtain an optimal dynamic regression model that satisfies the selected criterion.

The most universal optimality criteria used in the design of the experiment are the equivalent D- and G-optimality criteria, which simultaneously minimize the generalized variance of estimates for unknown model parameters and the maximum in terms of planning space variance of the output prediction according to the regression model [2-5].

When constructing dynamics models satisfying these criteria, it becomes necessary to formalize and solve additional problems:

- synthesis of optimal test signals [1, 4-11];
- technical implementation of optimal identification procedures in real time, which is especially important.

As a rule, DRMs have a nonlinearly parameterized structure. Therefore, their optimal identification is possible on the basis of the theory of sequential design of the experiment [2, 3], adapted to the problems of dynamics. Based on this theory, an original procedure of sequential D-optimal identification (DPPI) [1, 5-9] was developed, which allows us to study a wide class of linear and nonlinear dynamic objects described by input-output models and state spaces. Thus, the scientific novelty of this study lies in the development for the classical theory of experimental design and the development of methodology. Algorithmic and technical support of the original real-time control systems with optimal identifying studies of a wide class of dynamic objects.

2. Short description of DPPI

Let's Consider the class of dynamic objects described by regression models, represented in the form:

$$Y(n\Delta t) = \Phi(\vec{A},\vec{U}) + e(n\Delta t),$$

where $\vec{A}$ - is the vector of unknown parameters of the dynamic regression model; $Y(n\Delta t)$ - discrete value of the output of the object.

As the factors of planning discrete values of input signal are accepted:

*Fatuev Victor, Tula, Russia, +7(960) 597 74 46 vfatuev@inbox.ru

$$\vec{U}^T = (u((n-1)\Delta t), u((n-2)\Delta t), \ldots u((n-l)\Delta t))$$

DPPI is implemented in real time as follows.

At the time of starting the system, an arbitrary non-degenerate "seed" test signal is supplied to the input of the identifiable object through the actuator, the implementation of which determines the initial NLSM estimate of the vector of unknown model parameters and its covariance matrix. These data are used in planning the first stage of the procedure. When planning the next stage, the NLSM estimation and its covariance matrix are used, obtained from the results of all previous stages, including "seed" testing. The planning goal is to synthesize a part of the local D-optimal plan, which implies one or more planned exit measurements at the planned stage of the procedure. The choice of the number of planned measurements of the output is carried out on the basis of effectiveness criterion for the implementation of the procedure, taking into account the cost per unit of testing time, the cost of one measurement of the output, as well as the sensitivity function of the dispersion function of the DRM to estimates of unknown model parameters.

Each point of the plan corresponds to a piece of piecewise-constant testing signal, which must be implemented at the input of the object before the moment of the planned measurement of the output. The continuity of the test signal is ensured by "sewing" sections on γi clock cycles according to the algorithm:

$$u_i((n-l+j)\Delta t) = u_{i-1}((n-\gamma_i+j)\Delta t),\ i = 1,2,3,\ldots;\ j = 0,1,\ldots,(\gamma_i - 1),$$

where i– is the number of the sector.

If it is more efficient to plan several exit measurements at the stages of DPPI, then the synthesis of part of the local D-optimal plan implemented at the (N+1) - th stage is carried out according to the algorithm:

$$\vec{U}_{N+1,i} = \arg\max_{\vec{U}\in U_0} \nabla\Phi^T\left(\hat{\bar{A}}_N, \vec{U}\right)\operatorname{cov}\left(\hat{\bar{A}}_{N,i-1}\right)\nabla\Phi\left(\hat{\bar{A}}_N, \vec{U}\right),$$

$$\operatorname{cov}\left(\hat{\bar{A}}_{N,i}\right) = \left(I - \frac{\operatorname{cov}\left(\hat{\bar{A}}_{N,i-1}\right)\nabla\Phi\left(\hat{\bar{A}}_N, \vec{U}_{N+1,i}\right)\nabla\Phi^T\left(\hat{\bar{A}}_N, \vec{U}_{N+1,i}\right)}{\sigma_e^2 + \nabla\Phi^T\left(\hat{\bar{A}}_N, \vec{U}_{N+1,i}\right)\operatorname{cov}\left(\hat{\bar{A}}_{N,i-1}\right)\nabla\Phi\left(\hat{\bar{A}}_N, \vec{U}_{N+1,i}\right)}\right) \times$$
$$\times \operatorname{cov}\left(\hat{\bar{A}}_{N,i-1}\right),$$

where $\vec{U}_{N+1,i}$ - is i-th point of the part of the local D-optimum plan, defining (N+1)-th stage of DPPI; $cov(\hat{\bar{A}}_{N,i})$ - covariance matrix of vector $\hat{\bar{A}}_N$, if to specify it due to the results of i planned measurements of the output on (N+1)-th stage.

The implementation of DPPI ends at the stage in the planning of which the difference between the locally optimal plan synthesized in parts and the D-optimal one will become insignificant:

$$\left\| \max_{\vec{U}\in U_0} \nabla\Phi^T\left(\hat{\bar{A}}_N, \vec{U}\right)\operatorname{cov}\left(\hat{\bar{A}}_N\right)\nabla\Phi\left(\hat{\bar{A}}_N, \vec{U}\right) - \frac{q}{N}\sigma_e^2 \right\| < \varepsilon,$$

where q – is the number of unknown parameters of model; σ_e^2 - is dispersion of random output component; ε – a priori set small value.

Planning for the next stage begins from the moment the previous stage is completed. Therefore, it is relevant to minimize the planning time, which depends on the characteristics of the optimization and estimation algorithms used, as well as using the sequence of implementation for the plan points when constructing the test signal, taking into account the need for its level to remain constant over the planning interval.

A detailed algorithm for the implementation of DPPI is given in [1].

3. Description of the system architecture

A DPPI implementing system should consist of the following elements:

- Calculation Computer;
- programmable logic controller (PLC);
- sensors for measuring the output parameters of the identified object;
- actuators for realization of test signals.

The main functions implemented by the computer calculation computer [3-6,11]:

- planning of the DPPI stages;
- determination of discrete values for sections of minimum length testing signals and moments of planned output measurements.
- control of time instants for planned output measurements. At times when it is required to take data, the calculator issues a request for the programmable controller. It is supposed to use Ethernet as a communication port.

Functions implemented by the PLC:

- data collection from sensors of identifiable object output variables at planned time points;
- transfer of this data to the calculation computer for subsequent registration and processing. A diagram of the computing complex for solving problems of D-optimal sequential identification in real time is shown in [1]

The block diagram of a system of D-optimal sequential identification in real time is given in [1].

4. Calculation Computer

The software of the calculation computer implements 4 processes:

- the process of implementing the procedure of sequential D-optimal identification.
- the process of data requesting from the PLC;
- process of creating a control signal for actuator;
- scheduling process.

At each stage of the D-optimal identification procedure, the time moments of the change in the signal levels of the control action determining the piecewise-constant testing signal, and the

moments of requesting data from the output of the object are determined, which leads to the execution of processes 2 and 3.

The process of requesting data from the PLC occurs according to the following algorithm. At times when it is necessary to request data from the PLC, the software of calculation computer transfers the high execution priority to this process. A data request begin. After sending the request, the program will wait for a response of 100 ms. If no response is received, another request will be made. After 3 failed exchanges in a row, the program will end with an error. If the data was received, they will be written to shared memory [12] for process of sequential D-optimal identification. After completing this process, the software will block it until the next moment in time for requesting data from the PLC.

Exchange with the PLC will occur via Modbus TCP protocol. This is a modification of the MODBUS protocol for working in tcp/ip networks. Modbus is based on the Master-Slave architecture, where in this case the calculation computer will be the master, and the PLC will be the slave responding to requests.

The process of creating a control signal for actuators operates as follows. At times when it is necessary to change the testing signal, execution priority will be given to this process. A signal will be sent to the control device. After completing this process, the software will block it until the next moment in time of changing the level of the test signal.

During the operation of the software, it may turn out that the execution times of processes 2 and 3 may coincide, which may lead to the fact that one of these processes will not be completed in time. To eliminate this, process execution is spaced on different processor cores.

Fig. 1 shows how priorities are allocated and, accordingly, when the implementation of the corresponding process begins

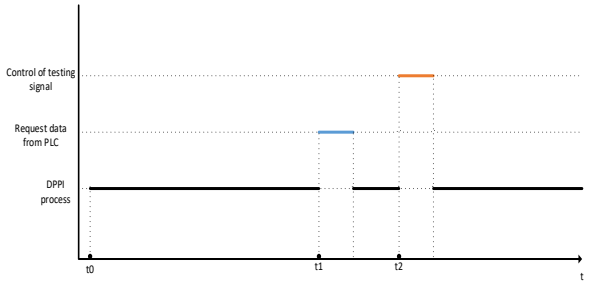


Figure 1: Prioritization of the processes of the calculation computer by time.

Fig. 1 shows that from time t0, which is the conditional start of the software, the high priority is always given to the process of implementing the sequential D-optimal identification procedure. The time instants t1 and t2 correspond to the time instants calculated during the planning phase of the DPPI.

Due to the fact that processes must be launched at certain points in time, it is required to implement another process that will monitor the execution time of the rest - the scheduler. The reason for its creation is that the RTOSs do not have their own software timers, and standard functions can lead to time fluctuations[12,13]. That why, own software timer needed to create. By using this timer, the scheduler will issue commands to transfer control to the processes described above.

5. Programmable logic controller

PLC is a device that works in cycles and performs two main functions

- inputs polling;
- execution of a user program that records the data received at the inputs and further transfers this data via Ethernet to the computer at the scheduled times.

To ensure high system reliability, it is required that the PLC respond to requests from the computer with high speed. This is achieved by the fact that:

- PLC provides a high frequency of data acquisition from sensors (estimated value of 100 Hz);
- The PLC, having received a request from the calculator at the scheduled time, instantly sends the last received data to the calculator.

PLCs are equipped with both discrete and analog inputs, which allows you to take data from the corresponding sensors without the use of additional equipment, such as analog-discrete converters. Data received from the PLC to the computer will be written to the data log file.

Most PLCs, in terms of receiving information, work with the MODBUS protocol based on the Master-Slave architecture [14-16], where the master is the PLC itself, while the slaves are surveyed sensors. The master initiates exchanges by sending requests. Such exchanges can be both individual and broadcast for all slave devices. The slave device responds to a request addressed to it. Upon receipt of a broadcast request, a response by slaves is not generated.

PLC connection to the object under study depends on the sensor, with which you can measure the output parameter of interest to us. Most often, these sensors have RS-485 and RS-232 communication lines, which will require analog ports and Modbus RTU protocol. Sensors operating over the Ethernet communication line are increasingly appearing. Such devices will require an Ethernet router, and the PLC will operate using the Modbus TCP protocol. In both cases, to receive data from sensors, PLC send request with the command 0x03 - registers reading [16].

In addition, PLC can be used as actuator, thanks to the presence of discrete inputs and outputs. The control is implemented as follows: the Calculation Computer sends a discrete output write command to the PLC - 0x05; when recording a discrete output, 2 values 0 and FF are used; Having accepted the command, the PLC, depending on the setting, will create a control action. The control action can be of two types:

- By analogue output. In this case, the numerical value of the test signal will be sent to the device for changing the input parameter. u1 for 0 and u2 for FF;
- By discrete output. In this case, the PLC sends a pre-programmed value received from the calculation computer to the actuator of the object

The algorithms for data request by the calculation computer and the operation of the PLC are given in [1].

The features of the technical implementation of DPPI for the identification of the reaction process of methanol and oxygen for the production of formaldehyde are considered in [1].

The developed system is universal and can be used for optimal identification of continuous technological processes in various industries.

Conflict of Interest

We declare there is no conflict of interest.

Acknowledgment

TSU (Tula state university), Tula, Russian Federation.

References

[1] Fatuev V.A., Mishin A.A. Realization of Optimal Identification Tasks for Dynamic Systems in Real Time Scale. 2019 8th Mediterranean Conference on Embedded Computing (MECO), 2019, pp. 537-540. DOI: 10.1109/MECO.2019.8760138

[2] Fedorov V.V. *Teoria optimalnogo experimenta* [Theory of optimal experiment], Moscow, Science, 1971, 312 p.

[3] Krug G.K., Fatuev V.A. Synthesis of D-optimal testing signal for one class of dynamic objects. *Avtomatika i telemekhanika* [Automation and Remote Control], 1974, no. 8, pp. 73-77. (In Russia)

[4] Krug G.K., Fatuev V.A. Synthesis of D-optimal testing signal for one class of dynamic objects. Avtomatika i telemekhanika [Automation and Remote Control], 1974, no. 8, pp. 73-77. (In Russia)

[5] Fatuev V.A. Optimalnaya identifikaciya i upravlenie dinamicheskimi sistemami. Ucheb. posobie [Optimal identification and management of dynamic systems. Allowance]. Tula, 2019. 120 p.

[6] Fatuev V.A., Khrapova A.G., Morozova A.N. Algorithmic ensuring realization of serial procedure D-optimal identification in real time. *Izvestiya TulGU. Tekhnicheskie nauki* [News of TSU. Technical science]. 2014, no. 11. part 2. pp. 612 – 620. (in Russia)

[7] Fatuev V.A., Mishin A.A. Control of the experiment with optimum identification of dynamic systems in real time. 2018 7th Mediterranean Conference on Embedded Computing (MECO). 2018. pp.1-4. DOI: 10.1109/MECO.2018.8406057

[8] Fatuev V.A., Mishin A.A. Algorithmic and technical ensuring for the implementation of the procedure for optimal identification of dynamic system in real time scale. *Izvestiya TulGU. Tekhnicheskie nauki* [News of TSU. Technical science]. 2018, no. 12. part. 2. pp. 315-323. (in Russia)

[9] Bahvalov L.A. On a problem of input signal synthesis on the statistical identification of linear dynamic systems. *Tekhnicheskaya kibernetika* [Technical Cybernetics]. 1974, no. 3. pp. 191-196. (in Russia)

[10] Boroduk V.P., Kirichenko A.V. Constructing the optimal input signal to identify multichannel dynamic objects. *Trudy MEI* [Works of MEI] 1975, no. 241. pp. 30 -36. (in Russia)

[11] Oslender D.M., Ridgeley J.R., Ringenberg J.D. *Upravlayushie programmi dlya mekhanicheskih sistem:obyektno-orientirovannoe proektirovanie sistem realnogo vremeni* [Control programs for mechanical systems: Object-oriented design of real-time systems], Moscow, 2004, 416 p.

[12] Burdonov I.B., Kosachev A.S., Ponomarenko V.N. Operacionnie sistemi realnogo vremeni [Real-time operating systems], Moscow, 2006, 98 p.

[13] Burdonov I.B., Kosachev A.S., Ponomarenko V.N. Operacionnie sistemi realnogo vremeni [Real-time operating systems], Moscow, 2006, 98 p.

[14] Minaev I.G., Sharapov V.M., Samoylenko D.G. *Programmiruemie logicheskie kontrolleri v avtomatizirovanih sistemah upravlenia* [Programmable logic controllers in automated control systems], Stavropol, 2010, 128 p.

[15] Parr E. *Programmiruemie kontrolleri: rukovodstvo dlya ingenera.* [Programmable Controllers: Engineer's Guide], Moscow, 2007, 516 p.

[16] Modicon modbus protocol reference guide - North Andover, 1996 – URL: http://modbus.org/docs/PI_MBUS_300.pdf

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 212-221 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Alternative Financing Model for Smart Cities Initiatives in Indonesia

Robertus Nugroho Perwiro Atmojo[1,*], Tota Pirdo Kasih[2], Yakob Utama Chandra[2]

[1]*Department of Information Systems, School of Information Systems, Bina Nusantara University, Jakarta, Indonesia*

[2]*Department of Industrial Engineering, Faculty of Engineering, Bina Nusantara University, Jakarta Code, Indonesia*

ARTICLE INFO

Article history:
Received: 24 October, 2019
Accepted: 25 November, 2019
Online: 22 January, 2020

Keywords:
Equity crowdfunding
Alternative financing
Indonesia Retail e-Commerce

ABSTRACT

Currently Indonesia is working actively in the implementation of the Smart Cities initiative program. However, to be able to apply the concept, there are three main elements that must be prepared as the basic foundation. These elements are structure, infrastructure, and superstructure. In this study we focused on the problem of the high cost of funding needed to finance these three elements. Therefore, we are of the opinion that the government must create new investment opportunities outside the stock exchange whose impacts can also be directly used for financing programs. In this study, we conducted an investigation of the interest of urban people to invest in online channels. The aim is to determine the decision-making behavior of the Indonesian people in non-consumptive capital expenditure activities when the government creates and publishes financing instruments on the market. The hope is to be able to create a national development financing strategy that is in accordance with the characteristics of the Indonesian people and right on target.

1. Introduction

During 2017 - until the present day, Indonesians' through several collaborative ministries (central government), local governments, the private sector, and the community have a high enthusiasm to improve and build a better nation. Led by the Ministry of Communication and Information, Ministry of Home Affairs, Presidential Staff Office, Ministry of Public Works and Public Housing, along with several other vital agencies. Currently, the nation is focusing on realizing a national development program in various aspects concerning the Smart Cities concept. The Smart Cities understood as a concept in which a region can manage its resources and support the living needs of its ecosystem independently. According to [1], [2], The Smart Cities concept includes six main aspects, namely Smart Governance, Smart Branding, Smart Economy, Smart Living, Smart Society, and Smart Environment. However, before being able to realize the Smart Cities program, it turns out that there are basic needs that must be met before being able to apply the concept.

Based on information obtained from the testimony of one of the compilers of the Guidance Book of Indonesia Smart City Master plan 2017, Ms. Dwi Elfrida (Person in Charge Indonesia Smart City Initiatives from Ministry of Information and Telecommunications). The implementation of the Smart City concept is analogous to a family that must already have a place to live before they can have a car for personal use. In other words, there are three main elements which are the pillars for implementing the Smart Cities concept in Indonesia, namely structure, infrastructure, and superstructure [2]. The structural elements include the development of human capital which is the agent and beneficiary of the Smart Cities concept. Within the structural elements are also discussed about budget resources and local government resources. The second element is infrastructure; infrastructure development includes three components, namely physical, digital and social infrastructure for the advantage of the community. The last element is the superstructure, in this section component such as regional policies and regulations, institutions, and procedures for implementing the Smart Cities concept must be accommodated.

The construction of three main elements are constrained by funding issues

Without having three adequate main elements, it is challenging for a region or even a country to implement the Smart Cities concept [3]. To be able to succeed in applying the concept, at least

* Robertus Nugroho Perwiro Atmojo, Email: robertus.atmojo@gmail.com

the people of a region must have a level of compliance and discipline towards proper legal regulations. Then to support innovation and economic growth, the potential must be supported by robust infrastructure and facilities. Finally, to be able to create a stable civilization on society, legal planners and implementers with integrity and reliability are needed. Unfortunately, in Indonesia, we are still struggling to resolve these crucial issues. The geographical conditions in the form of archipelagic countries and various types of local wisdom, make the development process of Indonesia have many challenges. Also, we suspect, one of the biggest challenges in the process is about infrastructure problems.

Infrastructure development is one of the essential elements that must be carried out by the state to make the nations' wheel of life continue to keep pace with the times. Without the availability of access to highways, railroads, seaports, airports, warehouses, optical fiber, radio towers, power plants, clean water treatment plants, and various other physical and digital infrastructures. So it is impossible for a country to continue to survive and develop in an era of increasingly intense global competition [4]. Unfortunately, infrastructure development is not as easy and cheap as imaginable. In Indonesia, to build the Trans Java toll road project along the 522 Kilometers, it is estimated that it requires an investment cost of IDR 17 Trillion (US$ 1.5 Billion), and even then only for the land acquisition budget [5]. Building nine dams in East Nusa Tenggara, West Nusa Tenggara, Banten, Central Java, Batam, and South Sulawesi, the government estimates that the project may require a national budget of IDR 3.85 Trillion (US$ 300 Million) [6]. Also, for the construction of the Trans Kalimantan railroad network infrastructure for 2400 Kilometers, it is estimated that the development requires an investment cost of IDR 22 Trillion (US$ 2 Billion) [7]. Regarding the example of the Trans Kalimantan infrastructure development project, it is regrettable that the project was one of fourteen lists of National Strategic Projects that had to be canceled by President Joko Widodo in 2018 [8].

Seek for alternative national development funding

The cancellation of the implementation of fourteen National Strategic Projects by President Joko Widodo made sense, considering the required investment costs were estimated to reach IDR 264 Trillion (US$ 20 Billion) [9]. Whereas at the same time, the nation still had to share its focus on the development of structural and superstructure elements. The Indonesian government's efforts to continue development also do not just stop. Issuance of investment instruments such as Saving Bonds Retail series 004 has been issued by the Ministry of Finance in 2018 to obtain a national financing capacity of IDR 1 Trillion (± US $ 100 Million) with coupon rates of 8.05% per annum [10]. In 2019, through the Ministry of Finance, the Indonesian government again issued Saving Bonds Retail series 005 for a national financing capacity of IDR 5 Trillion (± US$ 500 Million) with coupon rates of 8.15% per annum [11]. Government initiatives in issuing Saving Bonds indicate that for continuing national development, there are still many alternative financing sources needed. The good news from the issuance of the Retail Saving Bonds is quite exciting regarding the interest of the Indonesian middle class to contribute their capital toward national development. We assume that the community intentions formed are supported by trust building mechanisms that have been successfully campaigned by the government. Furthermore, the availability of a variety of instruments and channels for investment access both online and offline also helped facilitate community action to invest. Therefore, we suspect that now is the right moment for the government to conduct national revenue re-engineering. Specially to experiment in creating innovative instruments and investment access channels. In this study, we attempt to investigate the interest of urban people to invest in online channels. The aim is to determine the decision-making behavior of Indonesian people in non-consumptive capital expenditure activities. The results obtained from this study are expected to create social development financing opportunities that are following the characteristics of the Indonesian people.

2. Literature Review

A lot of new types of business offering profit sharing and business capital cooperation by financially based technology startups became our source of inspiration in this study. In Indonesia, the Go-Jek transportation ride-sharing startup that has adapted the Uber concept has successfully transformed into one of the financial technology startups that are quite important for the empowerment of local communities. Tokopedia and Bukalapak have fostered an entrepreneurial spirit for workers and retirees to continue developing their businesses. The Bareksa Investment Marketplace allows people to develop their capital not in conventional bank deposits, but on mutual fund exchanges. Bank Negara Indonesia also creates online securities products (with a zaisan and e-smart platform) that enable students to learn stock investments on the market with only less than US$ 10 deposits. More interestingly, in America, there is a crowdfunding-based online funding platform known as the Kickstarter. Where based on the company's claims, currently a total of successful projects funded have reached more than 250,000 projects — supported by around 15.5 million philanthropists, with funding turnover reaching US$ 4 billion.

2.1. Indonesian version of Kickstarter

Technological startups of crowdfunding-based business such as Kickstarter also exist in Indonesia. We know KitaBisa and Mekar as two examples of crowdfunding startups that collect enough philanthropists for business development, even for social financing. For KitaBisa, they claim that the donations have been collected reached IDR 512 Billion (US$ 45 Million). Indonesia also has iGrow, a crowd-investing based marketplace platform that brings farmers (or breeders), owners of agricultural land (or livestock), and investors synchronically. From the business model, iGrow claims to have empowered as many as 2200 farmers to manage more than 1,000 hectares of land, with yields of more than 500 tons only for peanut commodities (not including others) in Indonesia. Today's online channel-based crowdfunding and crowd-investment trends inspire us to test the intentions of the interest of the Indonesian people in making decisions to fund investment and charitable/social activities. To support the investigation, we chose retail-based e-Commerce channels as a testable activity.

2.2. E-Commerce channel

The e-Commerce channel is a dynamic, complex, and mysterious business ecosystem. There are many factors could

make a potential consumer just do online windows shopping or make a real decision when they visit online equity crowdfunding platform or an e-Marketplace. The factors can be influenced by technology or can be influenced by social factors. An excellent platform appearance, content information conformity, and ease of use are some fundamental aspects of technology that could attract customers to make purchases [12]–[14]. On the other hand, the ability of a retailer to provide protection for shopping and the quality of products offered is also an important aspect for motivating customer to purchase or to invest electronically [15]. Not only there, we assume there is also another factor that causes potential customers deciding to make a purchase, that factor is the quality of service [16], [17]. In our perspectives the world of retail e-Commerce, the good and poor quality of retailer services can be judged in two ways. The first aspect is the retailer's experience in fulfilling customer orders. The success of the retailers in completing the transaction process and meeting the needs of customer orders in online business can affect the reputation of the business entities both in the realm of online and conventional sphere. Because in previous study by [18], mentioned that the success of a retailer in delivering a satisfying shopping experience will build a sustainable online trust and it will impact on the seller electronic image. In the other study by [19] also proves that a good electronic image (e-Image) even can trigger consumer interest to pay for a product at much higher price than the suggested retail price (MSRP). The second aspect, we speculate that a retailer's ability to provide customer support/help is one of the keys to success that leads consumers to actualize their purchases. The previous research by [20], It is proof that the product recommendation provided by the retailer can make a potential consumer change product purchase decision. They could switch the purchase intentions of pre-planned products, into buying products recommended by the seller. Therefore, we believe that the presence of retailers in the online channel is also required. Through the electronic Word of Mouth [21], [22] and instant messaging technology for real-time online support assistance [23], [24]. Then, we think that Social Presence in the realm of online transactions is also important to be held.

2.3. Vendor reputation

Vavilis, Petković, & Zannone [25] defined reputation as a measure of online trust. The mechanism of reputation ranking of online business by consumers done in several approaches. As with the eBay auction site or Indonesian e-marketplaces like Tokopedia and Bukalapak, reputation ranking systems use star-scale systems (1 [bad] - 5 [good]) as a sign of the level of business reputation. In the Kaskus community forum, the user reputation rating system uses the colored beam scale (red: bad - Green: best) as the indicator. In a previous study by Gregg & Walczak [19], proving that high reputation ranking value and also a selection of the right business username can affect the level of customer trust. Where ultimately the level of trust formed correlates with the consumer's purchase decision [26]. Chen & Chou [27] defined the vendor's reputation as a reflection of the seller's ability to deliver the best service in the eyes of consumers.

The definition of vendor reputation in this research defined as vendor capability that manifests in recognition and public awards for producing the consistenly good performance. In other words, the positive digital footprint that vendors have because of good business history. Later it will be used as a measurement tool for prospective investors to establish business cooperation with them in the future (see Fig. 1). The measurement indicators of the supplier's reputation consist of:

(i) *The vendor's existence level*. The existence of suppliers in digital and physical channels must be easily recognizable and known to the general public. (ii) *The recognition/commendation level*. The amount of acknowledgment or award earned by a vendor from a community or third party appraisal institution. (iii) *Vendor ratings*. The amount of customer testimonials about the vendor on every products/services sold in internet.

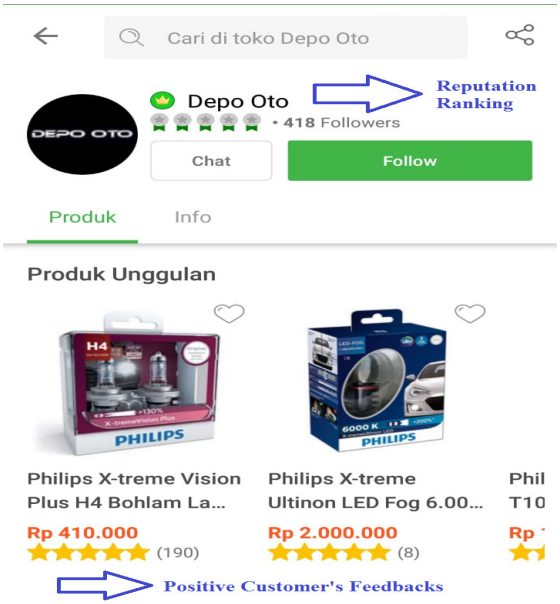


Figure 1. The example of vendor good reputation rating and positive customers' feedback (Depo Oto in Tokopedia.com)

2.4. Social presence

Social presence is a vendor/retailer manifestation in a business transaction mechanism. Social presence is difficult to create in online channels, especially if only rely on graphics or images [28]. However, as time passes and dynamic technological updates trigger the emergence of virtual reality concepts making social presence possible in online channels. The website is one of the key aspects in e-Commerce transaction. According to [22] social representation of online business players, psychologically possible presence by consumers through a particular medium, where the website or mobile site is the intended intermediary. Furthermore, to strengthen that perspective, it is appropriate that every online business person has a website or mobile media that is independent and not bound by third party intervention. Because, through independent ownership, the business has an unlimited discretion in serving consumers and perform regulation.

Further, ownership is personal as well as creates consumer perceptions of credibility, empathy, integrity, and benevolence (good impression good behavior). Ogonowski et al. [29] defined social presence as a perception of the personality, hospitality, and elements of the human sensitivity created in a website. In this

research the social presence is defined as a social approach given by the vendor to their consumer to reduce the service gap that occurs in the digital environment. In this case, the social presence measurement indicator divided into: (i) *Attention level*. Attention provided by the vendor in interacting and communicating with consumers online. (ii) *The vendor's ability level*. The ability of seller in delivering an atmosphere of online shopping experience that resembles the state of shopping traditionally to consumers. (iii) *Vendor politeness*. The ability of seller to express manners and ethics to their customers (see Fig. 2).



Figure. 2. The response from ZOTAC customer service agent that shows politeness and responsible business ethics for customer complaints in Newegg platform.

2.5. Online trust

Gefen [30] stated that consumer confidence in vendors is the basic necessity to obtain second, third, and onward purchase intentions. Alternatively, in other words, consumer loyalty is a manifesto of business success in building trustworthiness in the presence of consumers. Turilli et al. [18] also stated that trusts can not only occur in the real world (offline channels), but it can also form under online conditions. Therefore, the ability of vendors to build trust in the reality will reflect its integrity in digital channels and apply vice versa. Gefen [31] defined trusts in the context of e-Business and e-Commerce as the ability of consumers to take risks of uncertainty situations in transactions that can not be controlled by itself but still take action on certain expectations. According to [28], a definition of trust in the context of e-Business and e-Commerce is the ability of consumers to accept vulnerability owned by online vendors based on positive expectations and certain actions. Furthermore, Hwang & Lee [32] defined online trust as a consumer's ability to ignore uncertainty avoidance by using a reference to cultural norms prevailing in certain society as a guide to reducing risk of worry. In this case, allegedly consumers have a certain belief that things that are harmful to others are not appropriate to be done by someone (the trustee) because if it happened then the parties will receive certain social sanctions.

The ability of vendors to provide transparent, detailed, and accurate information is a trigger for consumer confidence online. Because in remote condition, consumers can not test the product directly, so it has become a duty for a seller to provide the online shopping experience that is not too much different from the traditional shopping experience (Fig. 3). This study defined online trust as a risk-taking mechanism conducted by consumers to depends on the performance capabilities of a vendor in conducting business transactions based on digital channels. And we agreed to have online trust consist of four dimensions, such as integrity, benevolence, ability, and predictability. The dimension of integrity measurement in this study consist of:

(i) *The level of confidence*. Consumers believe their familiar vendor will always be honest and will not cheat them or other customers. (ii) *Reliability level*. Consumers understand their online subscribed vendor will always have a reliable performance and full sense of responsibility for all actions that will accomplish. (iii) *The level of openness*. Online vendors must have transparency or openness in communicating with their customers so that consumers do not have to worry about cheating. Benevolence dimension (good behavior) in this study defined as the ability of the vendor to be able to act well and considerately in every action to the consumer. The measurement indicators of benevolences consist of (iv) *Confidence level*. Consumers believe that online subscribers always prioritize the interests of consumers compared with personal interests. (v) *Confidence level*. Consumers believe that their online subscribers will always act with consideration before performing their actions.

The ability dimension in this study defined as the level of competence possessed by the vendor in providing the best service for its customers. The measurement indicators of the ability dimension consist of: (vi) *The level of vendor competence*. Vendors must be able to demonstrate the competencies expected in providing shopping assistance to their consumers. (vii) *The rate of response speed*. It means, as a time of service that vendors can provide when consumers need help in the shopping process. (viii) *The vendor's troubleshooting competence*. A seller capability required in providing troubleshooting solutions in the event of a dispute with the consumer.

The dimension of predictability in this study defined as the ability of vendors to show performance patterns that can be predicted by consumers through a series of actions that follow the work procedures. The measurement of predictability indicators are as follows: (ix) *The rate of response speed*. The responsivity of vendors when answering questions and providing relevant advice needed by consumers. (x) *Expected level*. Consumers expect to always get an enjoyable online shopping experience from online vendors.

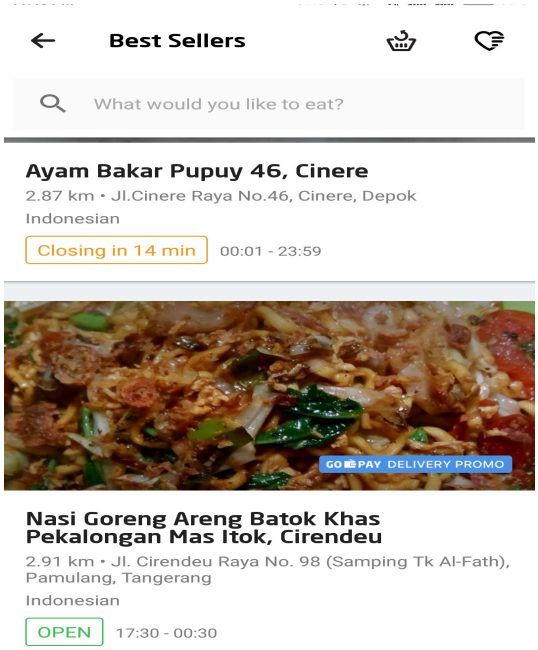


Figure 3. Nasi Goreng "Mas Itok" food vendor that are most preferred by Go-Food (Go-Jek) customers

2.6. Customer investment (purchase) intention

In this study, customer investment (purchase) intention is defined as the motivation of consumers to make investment decision towards equity offered (like in startup business cooperation opportunities). In other words, customer investment (purchase) intention in this research meant as a behavioral-based intention owned by a consumer to visit and make the realization of online purchases (or investment) from the desired business vendor [33]. The intention and desire are the last stages of the various considerations that consumers have before making a final purchase decision. Referring to Hong & Cho [34] measurement indicators of customer investment intention consist of (i) customer will return to visit the online store/vendor. (ii) Customer will make a purchase/investment from the online store/vendor within a period of two months. (iii) While new products are released in the market, then customer will prioritize purchases/investing from certain online stores/vendors. And (iv) for further purchases (or investment), customer will continue to visit and make actual decision from certain online vendor/retailer store.

3. Method

To support the analytic process in this research. We used convenience sampling as a method of data collection. The non-probabilistic sampling approach in this research is used to overcome the problem of dynamic and infinitive samples. For information, the sample used focuses on potential consumers who had shopping experience at online retail stores incorporated into the Indonesian e-Commerce Association (IdEA). As for some examples of these vendors are: included in the category of e-Marketplace (Tokopedia, Lazada, Bukalapak, BliBli) or providing marketplace platform (Go-Jek, and Grab) and also listed in the category of online retail stores (Traveloka, JD.id, Bhinneka).

The design of this study began with a scenario that questioning respondents intention to invest in online startups initiated by their familiar vendors. When their well-known vendors like Depo Oto (Fig. 1), Zotac (Fig. 2), or Fried Rice "Mas Itok," (**Fig. 3**) try to diversify their business in other fields. For example, Depo Oto, which is involved in the business of retailing automotive spare parts, will expand its business into agriculture development. Alternatively, Zotac, which is engaged in the Graphical Processing Unit, expands in the textile sector. And, if fried rice "Mas Itok" who started the business as a food seller innovating to build a semi-real estate property business. With consideration of all forms of administration and business legality guaranteed by Indonesian Government (like Otoritas Jasa Keuangan or The Ministry of Finance). Then we asked the respondents by several questions like:

"*Are the respondents interested on investing their capital in these startup companies?"* Considering they have had a business relationship (experienced) in the past. "*Can social capital that has been formed by both parties be developed to continue business relations and kinship between the two?".*

Some of these questions will then converted and tested through the relationship of several latent variables on the structural model. For the data collection technique, we used one shot cross-sectional and close-ended questionnaire based on a Likert scale with a range of values 1-5 (low-high). For data transformation process from qualitative scale to the quantitative scale used Method of Successive Interval. The goal of scale transformation helps to get a more weighted, stable, and objective assessment results. After the data collection process finished, the next phase is the process of data cleansing. The data cleansing process focused on using Mahalanobis Distance method to control outliers. The process of outliers handling is used to avoid the results of research from the use of biased data. In the last stages of data analysis, we used the method of Structural Equation Modeling based on Partial Least Squares (SEM-PLS). The purpose of using SEM-PLS in this research is to adjust the exploratory nature of research, where it requires a high accuracy measurement result by utilizing the nature of multiple regression analysis.

Refer to Fig. 4, the proposed hypotheses in this study consist of:

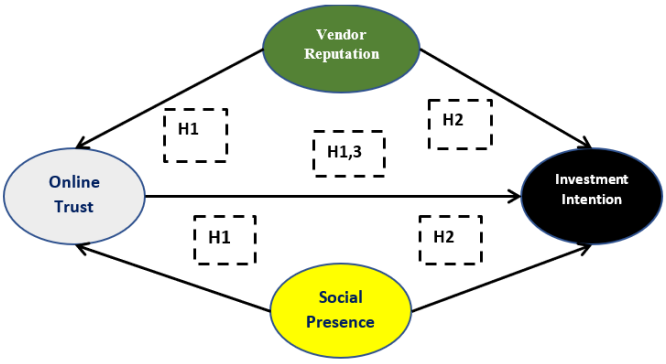


Figure 4. Relationship of structural model proposed in this research

H1: Vendor Reputation and Social Presence are positively correlated and have a significant direct effect on the establishment of Online Trust.

H2: Vendor Reputation and Social Presence are positively correlated and have a significant direct effect towards customer Investment Intention.

H3: Online Trust is positively correlated and has a significant direct effect towards customer Investment Intention.

4. Results and Suggestion

Out of the 350 distributed questionnaires results obtained as many as 323 respondents' answers are valid. The sample respondents are undergraduate students from a university in Jakarta majoring in information systems program. The average age of respondents is in the range 17-28 years. A total of 173 respondents are female, and the rest are male respondents. Based on the sample used, it assumed that participants are potential customers who have experience in the use of computers and smartphones. Moreover, it assumed they have a good understanding of internet usage [35]. Average monthly online spending from respondents ranged from IDR 200,000 - IDR 4,000,000 (US$ 15 – US$ 350). As for the data obtained, the most popular goods purchased by respondents when shopping online is a product of fashion and electronics. Fig. 5 shows the results of the first stage inner model testing using Smart-PLS software version 3. The first step of structural model testing can be seen in Fig. 5.

According to [36], we can say that the Online Trust moderately explains the creation of 40.5% variance in customer Investment Intention. Where, Vendor Reputation and Social Presence also have moderate explanation in Online Trust with the value of 42%. According to the path coefficient value, Online Trust has the strongest correlation towards customer Investment Intention with .457 value. Followed by Vendor Reputation → Online Trust with .366 value. Then Vendor Reputation → customer Investment Intention with .353 and Social Presence → Online Trust with .326 value. The weakest correlation coefficient indicated by Social Presence towards customer Investment Intention with -0.132 value.

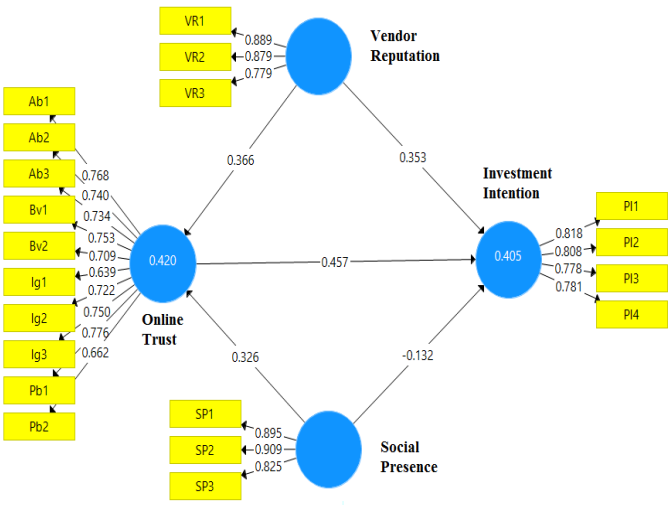


Figure 5. Inner model testing on structural model

In Table 1 it can be seen that all research variables have values of Cronbach α > 0.7 and rho_A > 0.4. In other words, we can say that all variables have met the required reliability standard in the measurement. In Table 1 it can also be seen that the value of Composite Reliability has exceeded 0.6 and the Average Variance Extracted has surpassed the value of 0.5. Thus, we consider all the measurement variables used to meet the criteria of structural reliability model.

Table 1. The Measurement of Reflective Outer Model

Latent Var.	Cronbach α	rho_A	Composite Realibility	Average Variance Extracted (AVE)
Online Trust	.90	.902	.918	.528
Investment Intention	.81	.817	.874	.635
Social Presence	.849	.849	.909	.769
Vendor Reputation	.807	.818	.887	.723

In Table 2 it can be noted that based on the calculation of Fornell-Larcker Criterion method all variables have met the required validity conditions. It indicated by the value of the intended variable (marked with the * sign) has exceeded the correlation value of other variables that reside in each column. For example, in the first column, the Online Trust variable has a diagonal value of .726.

That value exceeds the value of the correlation of other variables: Investment Intention (.593), Social Presence (.603), and Vendor Reputation (.612). In the second column also shows that the Investment Intention diagonal value has (.797) exceeded the Social Presence variable value (.410) and Vendor Reputation (.533). Thus, it can be said that based on the measurement of discriminant validity, then all research variables used have met the requirements of structural validity.

Table 2. Discriminant Validity Measurement

Latent Var.	Online Trust	Investment Intention	Social Presence	Vendor Reputation
Online Trust	**.726***	-	-	-
Investment Intention	.593	**.797***	-	-
Social Presence	.603	.410	**.877***	-
Vendor Reputation	.612	.533	.756	**.85***

4.1. Analysis after bootstrapping procedure

To perform hypothesis testing in SEM-PLS required bootstrapping procedure. The goal is to get more accurate measurements by combining core samples with sub-samples data. In this measurement, we use 5000 data as a sub-sample with two-tailed test mode, no-sign changes, complete bootstrapping, 2GB of memory usage, and 323 data as core samples.

Fig. 6 shows the results of inner model testing after going through the bootstrapping process. Table 3 shows the structural inner model test results after going through the bootstrapping procedure.

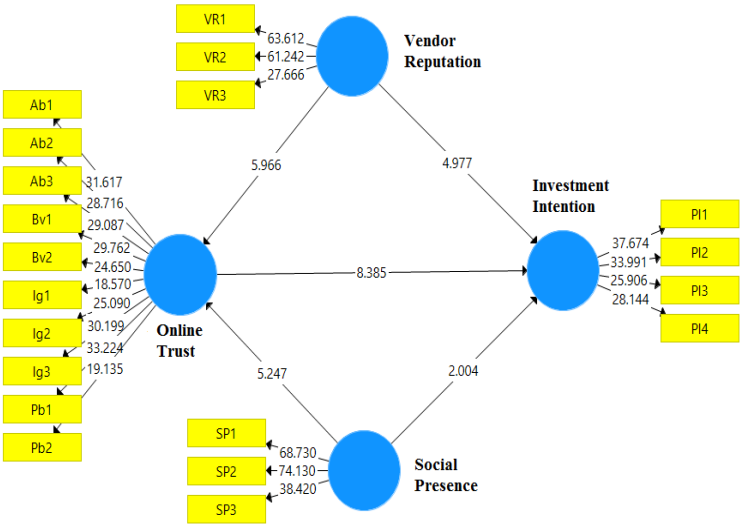


Figure 6. Inner model structure after bootstrapping

Table 3. Path Coeff. Values After Bootstrapping

Latent Var.	Original Sample	Sample Mean	StD	T-Stat	p-Values
Online Trust → Invest. Intention	.457	.459	.054	**8.385**	**.000**
Social Presence → Online Trust	.326	.327	.062	**5.247**	**.000**
Social Presence → Invest. Intention	-.132	-.132	.066	**2.004**	**.045**
Vendor Reputation → Online Trust	.366	.367	.061	**5.966**	**.000**
Vendor Reputation → Invest. Intention	.353	.353	.071	**4.977**	**.000**

The result of path coefficient measurement shown in Table 3 shows that in general, the correlation among the hypothesized variables has a good relationship. It is proofed by the correlation between Online Trust → Investment Intention through the value of T-Stat 8.385 > 1.96 (p-Val < .05). Furthermore, the value of Social Presence correlation → Online Trust has value T-Stat 5.247 > 1.96 (p-Val < .05). The correlation value of Social Presence → Investment Intention has T-Stat 2.004 > 1.96 (p-Val < .05). Correlation value of Vendor Reputation → Online Trust 5.966 > 1.96 (p-Val < .05). Moreover, the correlation between Vendor Reputation → Investment Intention has a T-Stat value of 4,997 > 1.96 (p-Val < .05). Where according to Wong [37] with a significance level of 5% then a good correlation should meet the requirements of T-Stat > 1.96 and p-Val < .05.

4.2. Hypotheses testing

In Table 4 we described the final result of structural model analysis. We use total effects as a measure of hypotheses testing.

Based on the result of total effects testing shown in Table 4. It can be concluded that:

H1: Vendor Reputation and Social Presence are positively correlated and have a significant direct effect on the establishment of Online Trust. (**H1 accepted**, p-Val < .05 & p-Val < .01).

Table 4. The Total Effects

Latent Var.	Original Sample	Sample Mean	StD	T-Stat	p-Values
Online Trust → Invest. Intention	.457	.459	.054	**8.385**	**.000****
Social Presence → Online Trust	.326	.327	.062	**5.247**	**.000****
Social Presence → Invest. Intention	.017	.018	.069	**.251**	**.801**
Vendor Reputation → Online Trust	.366	.367	.061	**5.966**	**.000****
Vendor Reputation → Invest. Intention	.52	.522	.070	**7.428**	**.000****

H2: Vendor Reputation and Social Presence are positively correlated and have a significant direct effect towards customer Investment Intention. (**H2 rejected**, Social Presence → Investment Intention has p-Val > .05).

H3: Online Trust is positively correlated and has a significant direct effect towards customer Investment Intention. (**H3 accepted**, p-Val < .05 & p-Val < .01).

From our measurement results, we found that Vendor Reputation and Social Presence variables have a strong correlation with the Online Trust it shown with the T-Stat > 1.96 and p-Values < .05 for both variables. Vendor Reputation and Social Presence together also have a significant influence on the establishment of Online Trust, it is shown with the value of p-Values < .05. Furthermore, the Online Trust Variables also have a strong correlation with customer Investment Intention variables; it has shown with the value of T-Stat > 1.96 and p-Values < .05. Similarly, the results of total effects measurement, Online Trust variable has a significant influence on the formation of customer Investment Intention with p-Values < .05.

What is interesting in this study is that we found the fact that although the Social Presence has a good correlation with customer Investment Intention (T-Stat [2,004] > 1.96 & p-Val [.045] < .05). However, the Social Presence variable does not have a significant effect on the formation of customer Investment Intention. It is shown with p-Val of total effects that exceed the 5% significance value (.801 > .05). We believe, the influence of Social Presence is not significant to the formation of customer Investment Intention caused by the sample respondents used is the people with a bachelor degree.

In other words, respondents with a decent level of education correlate with their level of understanding about the product and the use of a good e-Commerce retail platform as well. Therefore, allegedly that when they come to investing their capital towards products/services/equities online, respondents from academic societies prioritize vendor reputation compared to the sense of representation of the retailer. So we conclude that as long as the retailer can present a qualified product/service (worth to funded) and followed by a good reputation, then the consumer (especially from academic circles) will still make an investment, regardless of social presence of the retailer.

4.3. Opportunity for Indonesian government

The results of our investigation show that for university students, with a regular budget of IDR 200,000 - IDR 4,000,000 (US$ 15 - US $ 350) per month. The interest in funding vendors with excellent products and reputations has positive results. In other words, this is an opportunity for the government to take advantage of current possibilities regarding National Strategic Projects funding, especially infrastructure. We see opportunities for "government logos" as shown in **Fig. 7** will be an added value for the government when conducting investment campaigns for alternative funding. According to [38] Logos known as trust marks can be useful as accelerators for the creation of trust building mechanism on digital channels. Because with the existence of trust marks, we can say that the product/service/equity issued in the market has passed specific complex procedures and can also be used as an illustration of the reputation vendor.



Figure 7. Trust-Marks from the government

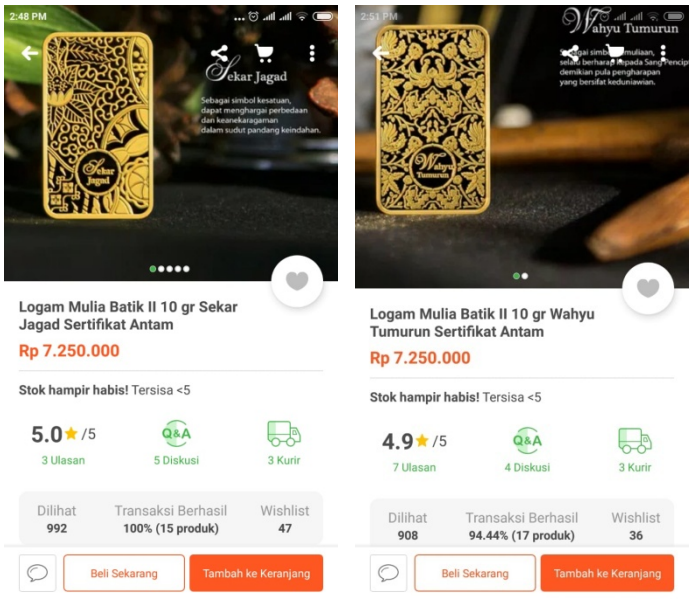


Fig. 8. Products from very good reputable vendor combined with arts are more expensive than regular products. Source: Toko Agung Cikini (Tokopedia's gold/jewelry retailer)

Government campaigns regarding investment programs for the community can be put into various manifesto. A simple instance is the issuance of precious metal instruments and jewelry that have high intrinsic value. As an illustration, our observation of some precious metal retailers on the Tokopedia platform shows findings where gold bar products with the themes of "*Sekar Jagad*" and "*Wahyu Tumurun*" batik (Fig. 8) sold at a margin of 10% more than the market price of similar products, and there are still buyers. (see Table 5 for product price comparison). It can be imagined if the central and regional governments in Indonesia collaborate with Antam (or other credible companies) to issue limited edition precious metals for infrastructure funding purposes. Of course, from the retail sales margin, the government has the opportunity to fund several projects without disrupting regional/central spending.

Table 5. Gold Bar prices in Indonesia market (price per 31st January 2019)

Certification Issuer	Denomination	Retail Price	Retailer
Antam (IDX ANTM)	1 Gram	IDR 693,000	PT. Pegadaian
Antam (IDX ANTM)	2 Gram	IDR 1,353,000	PT. Pegadaian
Antam (IDX ANTM)	5 Gram	IDR 3,280,000	PT. Pegadaian
Antam (IDX ANTM)	10 Gram	IDR 6,478,000	PT. Pegadaian
Antam (IDX ANTM)	25 Gram	IDR 16,093,000	PT. Pegadaian
Untung Bersama Sejahtera (UBS)	0.5 Gram	IDR 366,000	PT. Pegadaian
Untung Bersama Sejahtera (UBS)	1 Gram	IDR 659,000	PT. Pegadaian
Untung Bersama Sejahtera (UBS)	2 Gram	IDR 1,302,000	PT. Pegadaian
Untung Bersama Sejahtera (UBS)	5 Gram	IDR 3,188,000	PT. Pegadaian
Untung Bersama Sejahtera (UBS)	10 Gram	IDR 6,345,000	PT. Pegadaian
Untung Bersama Sejahtera (UBS)	25 Gram	IDR 15,816,000	PT. Pegadaian
Antam (IDX ANTM)	1 Gram	IDR 699,000	Indo Gold
Antam (IDX ANTM)	25 Gram	IDR 16,039,000	Indo Gold
Antam (IDX ANTM)	50 Gram	IDR 32,200,851	Indo Gold
Antam (IDX ANTM)	100 Gram	IDR 64,323,100	Indo Gold
Antam (IDX ANTM) (Sekar Jagad edition)	10 Gram	IDR 7,250,000	Toko Agung Cikini
Antam (IDX ANTM) (Wahyu Tumurun edition)	10 Gram	IDR 7,250,000	Toko Agung Cikini
Antam (IDX ANTM) (Sekar Jagad edition)	20 Gram	IDR 13,600,000	Toko Agung Cikini
King Halim (Shuang Xi edition)	5 Gram	IDR 3,170,000	Toko Agung Cikini
King Halim (Dragon edition)	5 Gram	IDR 3,170,000	Toko Agung Cikini
King Halim (Balinese Dancer edition)	5 Gram	IDR 3,150,000	Toko Agung Cikini
King Halim (Masjidil Haram edition)	10 Gram	IDR 6,190,000	Toko Agung Cikini

From Table 5, it can be noted that the price of gold bullion in the Indonesian market is influenced by the issuer institution which certifying metal authenticity. This fact strengthens the results of our investigation that reputation vendors have a higher bargaining power towards the market. Besides, the availability of goods supply factor is another determinant ascertaining the high selling

prices of commodities in the market. In this context, Antam's precious metals with the theme "Sekar Jagad" and "Wahyu Tumurun" were made as a limited edition work of art. Limited production quotas will result in a scarcity of products on the market, which also increases consumer interest in having them [39].

The U.S Geological Survey report supports this fact [40], Where in 2016 and 2017 for gold commodities, Indonesia managed to produce as many as 80 tons regularly, this was only evaluated by one commodity, according to a report from the Indonesian Ministry of Trade [41], in 2014 Indonesia managed to export 8mm size Sea-shells on the global market with a value of US$ 28 million. Thus there are still many other opportunities that Indonesia can do for the needs of national development funding.

4.4. Investment engineering

For national development, all efforts must be made. Investment engineering is a way that must be taken to deal with the limitations of physical assets provided by nature or to deal with secondary financing which does not need to be shouldered by the government itself. It is time for the government to expand the concept of economic sharing with the community. Whether it is investment sharing or benefit sharing. The Government of Indonesia, through one of the State-Owned Enterprises, PT Waskita Karya has attempted to carry out economic sharing with the business sector by selling several toll road segments for the capital of other toll road project funding [42]. It is perfect and needs to be appreciated. However, the targeted market is still specific; in other words, only giant capital companies can invest in toll road projects; the administrative process will also be complicated. Therefore, involving all levels of society by eliminating administrative problems such as "*minimum investment of US$ 100*" or "*return on investment can be obtained two years later*" will undoubtedly be a powerful way to boost up the source of capital. Indonesian people need to learn from the video game industry to follow the way they do financial engineering as is done by steam-powered platforms from Valve video games company in the United States. They realize that the business model in the game industry that produces games and then sells them off the market will undoubtedly be very dull and make the market become saturated quickly. The way that Valve does to overcome this problem is by creating a community market.

The concept of a community market is to allow end-users to trade game items they obtain from gamification results (such as reward points when winning matches or leveling up) in steam-powered marketplace platforms. The concept of the community market applied by Valve, in our opinion has been successful in making playing video games not dull anymore. A win for all concept, both for platform owners, developers, and end-users. Furthermore, a member in the Reddit community forum once predicted the question related to "*how much money is produced by Valve only from every game item traded in the community market?*". Although the question looks like jokes, the money obtained by Valve is estimated to be very serious. Throughout 2012-2013. with a 15% discount on every successful sale of game items, it is estimated that from ± 58 million transactions and trading volume of ± US$ 39 Million, Valve managed to obtain ± US$ 6 Million [43].

5. Conclusion

The results of our investigation showed that there is good interest from the public regarding the decision to invest online, especially those vendors who have known their reputation. This situation is also strengthened by the fact that the products sold by reputable vendors have the opportunity to carry out a pricing premium strategy (selling at higher prices because vendors have a bonafide business image). The investment (financial) engineering method has also been very diverse, even in the video games industry. So we conclude that now is the exact moment for the Indonesian government to make creative breakthroughs to create national project funding opportunities. Because the available market for breakthroughs is not only from the local community but globally. If the government is willing to do that, then we believe that the Smart Cities concept will be successfully implemented in Indonesia and will bring prosperity to all people.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] R. Giffinger, "Smart cities Ranking of European medium-sized cities," October, vol. 16, no. October, pp. 13–18, 2007.

[2] I. D. G. of I. Application, Buku Panduan Penyusunan Masterplan Smart City 2017. Gerakan Menuju 100 Smart City. Jakarta, Indonesia, 2017.

[3] R. Mahesa, G. Yudoko, and Y. Anggoro, "Platform Ecosystems for Indonesia Smart Cities," in 2018 International Conference on Computer, Control, Informatics and its Applications: Recent Challenges in Machine Learning for Computing Applications, IC3INA 2018 - Proceeding, 2019, pp. 34–39.

[4] R. Mahesa, G. Yudoko, and Y. Anggoro, "Dataset on the sustainable smart city development in Indonesia," Data Br., vol. 25, pp. 1–16, 2019.

[5] A. Ramadhiani, "Pemerintah Habiskan Rp 17,27 Triliun demi Lahan Tol Trans Jawa," Properti Kompas.com, Jul-2017.

[6] E. Simorangkir, "9 Proyek Bendungan Jokowi Rp 3,85 T akan Rampung Tahun Ini," detikFinance, Jan-2018.

[7] M. Idris, "Berapa Biaya Bangun 2.428 Km Rel Baru di Kalimantan?," Detik Finance, May-2017.

[8] H. Kusuma, "Ini Daftar Lengkap 14 Proyek yang Dicoret Jokowi dari PSN," Apr-2018.

[9] T. Hamdani, "Saat Jokowi Coret 14 Proyek Strategis Nasional Senilai Rp 264 T," Dec-2018.

[10] M. Wareza, "SBR 004 Beri Bunga 8,05%, Kemenkeu Cuma Incar Dana Rp 1 T," CNBC Indonesia Investment, Aug-2018.

[11] H. K. Dewi, "Kuota Nasional SBR005 Tinggal Rp1,47 Triliun Lagi, Ayo Segera Pesan," Bareksa News Funds, 2019. .

[12] D. Cyr, "Website design, trust and culture: An eight country investigation," Electron. Commer. Res. Appl., vol. 12, no. 6, pp. 373–385, 2013.

[13] R. N. P. Atmojo, A. D. Cahyani, B. S. Abbas, B. Pardamean, Anindito, and I. D. Manulang, "Design of single user decision support system model based on fuzzy simple additive weighting algorithm to reduce consumer confusion problems in smartphone purchases," Appl. Math. Sci., vol. 8, no. 13–16, pp. 717–732, 2014.

[14] Y. Lie, R. N. P. Atmojo, H. H. Muljo, S. Purnamasari, and B. Pardamean, "A comparison of customer preference towards two different types of e-Tailing channel," in Proceedings of 2017 International Conference on Information Management and Technology, ICIMTech 2017, 2018, vol. 2018-Janua, pp. 171–176.

[15] R. N. P. Atmojo et al., "Rekening Bersama (RekBer) The Indonesian Escrow Service of Money," Int. J. Multimed. Ubiquitous Eng., vol. 11, no. 3, pp. 399–410, 2016.

[16] A. O. Ajayi, G. A. Aderounmu, H. A. Soriyan, and A. David, "An intelligent Quality of Service brokering model for e-commerce," Expert Syst. Appl., vol. 37, no. 1, pp. 816–823, 2010.

[17] V. U. Tjhin, R. Tavakoli, and R. N. P. Atmojo, "The Determinants Affecting E-Loyalty: Hospitality Industry in Indonesia," in 2016 2nd International Conference on Science in Information Technology (ICSITech), 2016, pp.

131–136.
[18] M. Turilli, A. Vaccaro, and M. Taddeo, “The Case of Online Trust,” Knowledge, Technology & Policy, pp. 1–13, 2010.
[19] D. G. Gregg and S. Walczak, “Dressing Your Online Auction Business For Success: An Experiment Comparing Two Ebay Businesses,” MIS Q., vol. 32, no. 3, pp. 653–670, 2008.
[20] R. N. P. Atmojo, Anindito, B. Pardamean, B. S. Abbas, A. D. Cahyani, and I. D. Manulang, “Fuzzy simple additive weighting based, decision support system application for alternative confusion reduction strategy in smartphone purchases,” Am. J. Appl. Sci., vol. 11, no. 4, pp. 666–680, 2014.
[21] C. W. Yoo, G. L. Sanders, and J. Moon, “Exploring the effect of e- WOM participation on e-Loyalty in e-commerce,” Decis. Support Syst., vol. 55, no. 3, pp. 669–678, 2013.
[22] E. Toufaily, N. Souiden, and R. Ladhari, “Consumer trust toward retail websites: Comparison between pure click and click-and-brick retailers,” J. Retail. Consum. Serv., vol. 20, no. 6, pp. 538–548, 2013.
[23] R. N. P. Atmojo et al., “Research Plan Development Concerning E-Image Impact Towards Online Purchase Intention And Premium Pricing Strategies In Indonesia Community Based Online Market,” in 2015 International Conference on Science in Information Technology (ICSITech), 2015, pp. 1–6.
[24] R. N. P. Atmojo and A. D. Cahyani, “Recommendations for Development of Online Bookstore in Indonesia: a Review From Writer’s Perspective,” J. Comput. Sci., vol. 10, no. 8, pp. 1522–1530, 2014.
[25] S. Vavilis, M. Petkovi?, and N. Zannone, “A reference model for reputation systems,” Decis. Support Syst., vol. 61, no. 1, pp. 147–154, 2014.
[26] L. Liu, X. Zhao, H. Wang, W. Song, and C. Du, “Research on Identification Method of Anonymous Fake Reviews in E-commerce,” TELKOMNIKA, vol. 14, no. 4, pp. 1510–1520, 2016.
[27] Y.-T. Chen and T.-Y. Chou, “Exploring the continuance intentions of consumers for B2C online shopping Perspectives of fairness and trust,” Online Inf. Rev., vol. 36, no. 1, pp. 104–125, 2014.
[28] D. Gefen and D. W. Straub, “Consumer trust in B2C e-Commerce and the importance of social presence: Experiments in e-Products and e-Services,” Omega, vol. 32, no. 6, pp. 407–424, 2004.
[29] A. Ogonowski, A. Montandon, E. Botha, and M. Reyneke, “Should new online stores invest in social presence elements? The effect of social presence on initial trust formation,” J. Retail. Consum. Serv., vol. 21, no. 4, pp. 482–491, 2014.
[30] D. Gefen, “Customer Loyalty in E-Commerce,” J. Assoc. Inf. Syst., vol. 3, no. 1, pp. 27–51, 2002.
[31] D. Gefen, “E-commerce: the role of familiarity and trust,” Omega, vol. 28, no. 6, pp. 725–737, 2000.
[32] Y. Hwang and K. C. Lee, “Investigating the moderating role of uncertainty avoidance cultural values on multidimensional online trust,” Inf. Manag., vol. 49, no. 3–4, pp. 171–176, 2012.
[33] H. Xu, R. Zhang, C. Lin, and W. Gan, “Construction of E-commerce Recommendation System Based on Semantic Annotation of Ontology and User Preference,” TELKOMNIKA Indones. J. Electr. Eng., vol. 12, no. 3, pp. 2028–2035, 2014.
[34] I. B. Hong and H. Cho, “The impact of consumer trust on attitudinal loyalty and purchase intentions in B2C e-marketplaces: Intermediary trust vs. seller trust,” Int. J. Inf. Manage., vol. 31, no. 5, pp. 469–479, 2011.
[35] E. Symeonaki, M. Papoutsidakis, D. Tseles, and M. Sigala, “Post-implementation evaluation of a University Management Information System (UMIS),” in Proceedings - 2016 3rd International Conference on Mathematics and Computers in Sciences and in Industry, MCSI 2016, 2017, pp. 14–19.
[36] W. W. Chin, “The Partial Least Square Approach to Structural Equation Modeling,” in Modern Methods for Business Research, 1998, pp. 295–336.
[37] K. K. Wong, “Partial Least Squares Structural Equation Modeling (PLS-SEM) Techniques Using SmartPLS,” 2013.
[38] W. Wandoko, B. S. Abbas, D. Budiastuti, and R. Kosala, “Online trust building through third party trust transfer and third party protection,” J. Phys. Conf. Ser., vol. 801, pp. 1–8, 2017.
[39] S. Worchel, J. Lee, and A. Adewole, “Effects of supply and demand on ratings of object value,” J. Pers. Soc. Psychol., vol. 32, no. 5, pp. 906–914, 1975.
[40] U. S. G. Survey, “Mineral Commodity Summaries 2018,” 2018.
[41] T. M. of T. R. of Indonesia, “Indonesian South Sea Pearls,” 2016.
[42] A. Basith, “Waskita Karya jual 18 ruas tol untuk menutupi utangnya,” Kontan Industri, 2018.
[43] Reddit, “Valve’s Cash Cow - The Steam Community Market.” 2013.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 2, 215-224 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Enhancing an SDN Architecture with DoS Attack Detection Mechanisms

Ranyelson Neres Carvalho, Lucas R. Costa*, Jacir Luiz Bordim, Eduardo Adilo Pelinson Alchieri

Department of Computer Science - Institute of Exact Sciences - University of Brasília - 70910-900 - Brazil

ARTICLE INFO

Article history:
Received: 03 December, 2019
Accepted: 20 February, 2020
Online: 25 March, 2020

Keywords:
Software Defined Network (SDN)
Denial of service (DoS)
Attack identification
OpenFlow
Entropy
Chi-square

ABSTRACT

A Software Defined Network (SDN) architecture is characterized by decoupling the data plane and control plane. This feature enables the establishment of a programmable environment in which the control plane acts under the data plane, managing and configuring the network over a standard protocol, such as OpenFlow. Although there are numerous benefits to the SDN architecture, security is still a matter of concern, as the decoupling expands the attack surface on the network. Denial of service (DoS) attacks is one of the major challenges for the SDN architecture, mainly due to the vulnerabilities existing between the control and data planes. This paper proposes an enhancement to the SDN architecture enhanced with DoS attack detection mechanisms. Two techniques have been evaluated to identify of DoS attacks: entropy and chi-square. Both techniques use OpenFlow switches statistics so that it is possible to distinguish benign traffic from spurious traffic. Experimental results show that entropy and chi-square present similar results in terms of spurious flow detection accuracy. However, on average, chi-square requires about 14.89% *fewer packets to detect the attack when compared to entropy.*

1 Introduction

Software Defined Networks (SDN) is an upcoming network architecture that provides a more simplified and flexible way to manage a network. One of the main aspects of a SDN network is decoupling the data plane from the control plane. When this architecture is employed, the control plane is responsible for managing the entire network through operational decisions, coordinating and adding new functionalities. While the data plane is classified as the network infrastructure and its tasks are routing, dropping packets on the network, as well as transmitting information regarding packets to the control plane The communication between the two planes (control and data plane) is carried out through a protocol. The most widely used protocol is OpenFlow, which is characterized as a generic method of communication between the controller (control plane) and switches (data plane) [1]. Despite the numerous benefits provided by an SDN architecture, network security is still a matter of concern [2]. Decoupling the two planes presents new challenges to network security, since in conventional networks, security mechanisms are often defined through a combination of configurations between both planes [3].

Denial of Service (DoS) attacks are one of the major challenges for a SDN network, mainly due to the vulnerabilities present in the communication between the two planes [4]. In this architecture, the main goals of a DoS attack are to exhaust the controller's processing capacity or reduce the bandwidth of an application, which is usually accomplished by generating large volumes of traffic through multiple hosts with spoofed addresses [5].

DoS attack detection has been broadly researched in traditional networks, where most of the proposed works employ knowledge-based methods, as can be seen in [6, 7, 8]. However, the use of statistical methods or machine learning also proves to be effective for detecting a DoS attack. Such methods are commonly used in the literature to identify deviations in traffic behavior patterns, presenting high accuracy for this purpose. In [9], for example, the authors suggested a scheme based on machine learning to detect DoS attacks in conventional networks, where they propose a neural network designed to classify legitimate from spurious traffic.

In contrast to the aforementioned works, this article proposes a DoS attack detection mechanism in SDN. In addition, this paper proposes an enhancement to the SDN architecture with a mechanism that works alongside the controller to detect DoS attacks. Two techniques are evaluated to detect DoS attacks: **entropy** [10] and **chi-square** [11]. In the OpenFlow protocol, the controller can collect information about the switches to calculate statistics. These values are used to differentiate spurious traffic from legitimate traffic. The main contributions of this paper are a mechanism that allows the definition of an adaptive detection threshold that may reflect optimal network behavior and the use of different OpenFlow protocol header fields to characterize DoS traffic. Among its features, the

*Corresponding Author: Lucas Rodrigues Costa, Campus Darcy Ribeiro. Asa Norte, Braslia, DF - 70910-900 - Brazil, lucasrc.rodri@gmail.com

proposed mechanism encompasses the following characteristics:

1. the implementation resides at the control plane SDN, allowing it to obtain data from the whole network with greater precision, instead of a specific switch as used in a conventional network;

2. development of a mechanism to detect DoS attacks using OpenFlow switches statistics, thus enabling the use of different OpenFlow protocol header fields to characterize DoS traffic.

Experimental evaluation of the proposed solutions indicates that the entropy-based mechanism is slightly more accurate in identifying spurious flow at the cost of requiring a larger number of packets to detect attack compared to the chi-square-based mechanism.

The remainder of this paper is structured as follows. Section 2 introduces the SDN concepts. Section 3 presents how SDN DoS attacks work and how chi-square-based and entropy-based attack detection mechanisms work. Section 4 presents the anomaly detection techniques (entropy and chi-square) and how they can be used to detect DoS attacks. In Section 5 presents a brief overview of the related works. Section 6 introduces the proposed detection method applied to an SDN. Finally, Section 7 presents the simulation results, followed by the conclusion in Section 8.

2 Software Defined Networks and OpenFlow

The OpenFlow protocol is known as one of the first standards for Software Defined Networks (SDN), it enables communication between data planes and switch. It also allows the management of the query table used by the hardware to specify the next step of each received packet. There is a logical division between the data and control layers (a.k.a. planes), which establishes a programmable environment [1]. The planes that make up the SDN architecture are depicted in Figure 1.

The application layer consists of SDN applications for high-level management of network resources and support of the control layer below. The north interface is used for establishing the communication between the application layer and the control layer [12]. The control layer is in charge of programming and managing the data plane. It is composed of controllers that use the data obtained by the data layer and establishes all the procedures and actions to be performed by the network infrastructure [13].

The infrastructure layer (data plane) is made up of forwarding devices (switches and routers). This layer is responsible for local information monitoring and statistical collection [1]. The south interface is responsible for the communication between data layer and controller [14]. The best-known interface is the OpenFlow protocol, which allows the controller to observe the flow table of the switches, in order to implement actions through a standardized interface.

The protocol architecture is composed of: (*i*) OpenFlow switches that form the data plane; (*ii*) the control plane, which includes one or more controllers; and (*iii*) a secure communication channel, which links the switches to the control layer. An OpenFlow switch consists of a flow table and a secure communication channel. The flow table is formed by a set of input flows in which each input has an associated behavior, such as discarding, forwarding and sending the packet to the controller. This table is formed by actions, rules and counters. The rules are constituted from one or more header fields and correlated to a set of actions, which specify how the packets will be processed. The counters store statistics for switch flows, such as bytes received and transmitted, number of packets and the elapsed time since the flow was inserted into the switch. This feature allows the collection of statistics that can be used for specific purposes, such as detecting network anomalies, estimating the delay and so on. The collection of statistics can be achieved with a flow or a group of flows. The secure communication channel is the path used to transport information between the controller and the switch and vice versa [15].

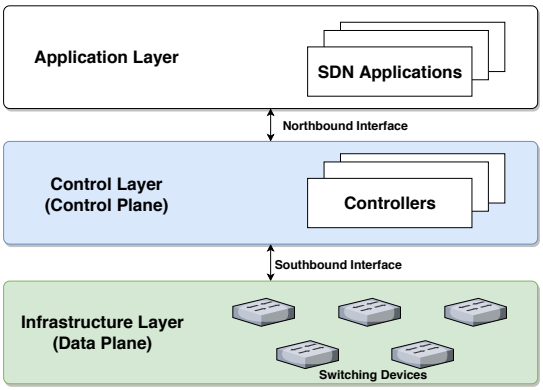


Figure 1: SDN architecture.

The packets processed by the OpenFlow switch are compared with the flow table and if a corresponding entry is found, the respective action will be taken. Otherwise, the packet will be forwarded to the control layer through the secure communication channel and the switch will wait for instructions from the controller. When receiving a packet, the controller performs the action that was previously established in its configuration for this type of flow. For example, remove or add new information from the flow table and apply the rule to the switches.

3 Denial of Service in SDN

A denial of service (DoS) attack on an SDN network can target both the data plane and the control plane [4]. These attacks are often known for generating excessive amounts of malicious traffic over a short period of time, in order to exhaust the resources of a victim or an infrastructure. DoS attacks are characterized by a high volume of traffic that floods the network, reducing or preventing a victim from being accessed [5]. In the data plane, this attack can deplete the flow table entries of the switches, through false streams generated by the attacker using the same source address [4]. In the control plane, the attack can cause a bottleneck in the communication channel between the controller and the switch [4]. In this situation, the attacker can exhaust the controller's processing efficiency by forwarding packets

with different headers that do not correspond to any flow already defined in the switch. This will force the switch to send control packets to the controller, since there is no rule to forward that malicious packet. DoS attacks can be even more complex when they are orchestrated in a distributed way and are known as Distributed Denial of Service (DDoS) attacks. These attacks can be even more detrimental to the network, while also increasing the complexity and the cost to defend the victim.

DoS attacks can be classified into two main groups: volumetric attacks and non-volumetric attacks (low volume or low rate) [16]. The latter covers attacks in which specific vulnerabilities of the communication protocol are exploited, causing the exhaustion of some of the victim's resources. The former, on the other hand, includes attacks where large volumes of traffic flood the victim. It aims to exhaust the victim's processing capacity or bandwidth preventing legitimate requests to reach the victim. Volumetric attacks can be further divided into direct and reflexive attacks. Direct attacks send traffic directly to the victim while reflection attacks use intermediate hosts (reflectors) to flood the victim. In this work, we restrict to SYN Flood-type attacks, characterized as a direct-volumetric attack.

The SYN Flood attack is based on sending a large number of TCP connection requests to the server, the SYN segments are sent with different source addresses, that are usually spoofed. For each SYN sent, the server must respond with a SYN-ACK segment, waiting for a confirmation message (ACK) to configure the connection. Since the source address of the SYN packet is usually false, the ACK response will never arrive because the SYN-ACK was sent to a nonexistent address. Causing TCP connections to remain in a semi-open state (SYN-RECV), where the memory resources are not released. With a large volume of false SYN segments sent, the server's TCP connection queue is flooded, causing the server to eventually crash. This, in turn results in a state in which the server cannot respond to any requests from TCP clients [17].

4 Anomaly Detection Techniques

Anomaly detection techniques help in identifying patterns outside the expected behavior. They can be performed through collection, filtering, and data processing.

Detection mechanisms play a key role in the DoS attack detection scenario. Beitollahi et al. [18] classifies the detection mechanisms based on the algorithm used for the detection. According to this classification, there are four groups: (*i*) statistical techniques; (*ii*) wavelets analysis; (*iii*) point of sequential change; and (*iv*) machine learning techniques. In statistical techniques, the profile of normal traffic flow is provided by the header information in the packet, which is then compared to the spurious traffic. Wavelet analysis techniques describe an input signal in terms of spectral components, which are representations of amplitudes or intensities. At the point of sequential change, filtered traffic is treated as a time series, i.e. a collection of observations made sequentially over time. Detection of a change in time is provided by evaluating all past data, where one must distinguish between two processes states: out of control or under control. Finally, machine learning techniques are capable of transforming data into intelligent actions, allowing decision-making to be performed automatically. This work focuses on statistical techniques.

Statistical techniques analyze various properties of specific fields in packet headers during normal traffic conditions. Then, by using statistical properties, the system creates a reference model for normal traffic. When the statistical properties calculated for the traffic show differences from the reference model, it is possible to observe traffic conditions that could indicate an attack. This kind of technique uses two basic components for detection: an observation window and a detection threshold. The size of the observation window can be based on the length of time or the number of packets to be analyzed. The threshold is used to capture variations that go beyond predetermined limits. Some statistical methods, such as chi-square and entropy, can report such type of information [6] [19].

4.1 *Chi-square (a.k.a. χ^2)*

Chi-square is a measure that portrays the difference between two probability distributions. The basic principle of this method is to compare proportions, that is, the possible divergences between the frequencies observed and expected for a certain event [20]. The equation is defined as follows:

$$\chi^2 = \sum_{j=1}^{m} \left[\frac{(o_j - p_j)^2}{p_j} \right], \tag{1}$$

where m is the total number of occurrences observed in a window w_i, o_j is the observed frequency for each class and p_j is the expected frequency for that class. When the frequencies observed are very close to the expected value, the chi-square value is small. Otherwise, chi-square will assume high values.

4.2 *Entropy*

Shannon entropy [10] is a well-known and valuable concept in information theory. It measures the probability of an event occurring concerning the total number of events, through which it is possible to describe the degree of dispersion or concentration of a distribution. The equation is defined as follows:

$$H = -\sum_{j=1}^{m} p_j \log p_j, \tag{2}$$

where m is the number of different occurrences in an observation window w_i and p_j is the probability associated with each occurrence j in the window. In other words, p_j is the occurrence frequency of every single item divided by the total number of items. The result of this calculation varies between 0 and $\log m$ [21]. Higher values indicates more disperse data in the distribution. Lower value indicates maximum concentration in the distribution. In other words, if the entropy is low then the randomness of the distribution is also low and vice versa.

5 Related Works

Research on the use of statistical methods to detect DoS attacks has grown over the years [20]. Statistical methods such as chi-square and entropy are great representatives of this class, as they present

excellent results for this purpose in traditional networks [22], [23], [7], [21].

One of the major problems with using such detection mechanisms in traditional networks is that each edge switch has only local network traffic, so it is not possible to have a global network scanning scenario. Therefore there is no mechanism that can provide a network overview. Such a limitation can make it difficult to detect an attack when it is distributed throughout the network. In addition, there is a relatively high cost of deploying the detection mechanism on each edge device that make up the network, which may become unfeasible depending on the infrastructure.

In the SDN context, there are some works that proposed the use of statistical analysis to evaluate possible changes in traffic from a DoS attack. In [24], the authors suggested a detection mechanism that uses data plane entropy from a programmable switch. The authors in [25] proposed the use of entropy in the control plane, taking advantage of the controller overview to detect traffic changes that may characterize the DoS attack. Although the mentioned works use entropy, they do not present the volume of packets necessary for the detection of the attack and other methods are not explored in the presented proposals.

This work proposes a solution for detecting DoS attacks in SDN that takes advantage of the controller's network traffic knowledge. That is, unlike traditional network approaches where each switch implements local strategies, we take advantage of the controller overview to detect anomalies. To be more specific, this paper implements and evaluates statistical DoS detection mechanisms, entropy and chi-square, which are incorporated at the controller. The details of the proposed solution are presented and evaluated in the subsequent sections.

6 Proposed solution to detect DoS attacks in SDN

In the SDN architecture, traffic flow monitoring and the execution of pre-established actions are performed by the control plane. The controller operates with the OpenFlow switches, exchanging information and statistical data via specific communication channels [3]. This data includes the number of bytes, the number of packets and the elapsed time that the flow was established on the switch. These can be forwarded to the controller in an organized way, through a timer, which requests the current set of data from the switch for a group of flows or a certain flow. All input streams are managed by flow tables in the switch. If a stream is not configured in its flow table, the switch sends an asynchronous message named *packet_in* to the controller. Otherwise, the stream follows the behavior determined by its flow table. This message is sent whenever there is no corresponding flow entry in the switch flow table. Thus, the controller will add a flow rule or eliminate flows from the flow table.

The use of spurious IP addresses is common in DoS attacks. In the course of an SYN Flood DoS attack, all spoofed source IP addresses target the victim's IP. In an SDN, there will be no entries in the flow tables for spoofed addresses, so the spurious flows will be forwarded to the controller via *packet_in*. If the packet rate is too high, such as in a DoS attack, the controller's resources will start to run out. In this situation, a high rate of spoofed addresses can overload the controller. The consequence of this will be the flooding of the communication channel between the data and the control plane, making it impossible for legitimate streams to be routed to their destination. Considering that it is possible to identify a new stream in the network and that during a DoS attack, this stream will have non-standard characteristics of the normal traffic. Then it is possible to measure this divergence through techniques that detect anomalies like chi-square and entropy. These techniques can identify changes in traffic that can be characterized as a DoS attack.

To develop the DoS attack detection mechanism using statistical analysis techniques it is necessary to define four components: observed field, observation window, statistical method, and the detection threshold.

6.1 Observed field

The available packet header fields to be analyzed include source IP address, destination IP address, source port, destination port, packet size, and etc. It is possible to extract packet header information at runtime. Each strategy should define the field to be analyzed that allows characterizing spurious traffic from legitimate traffic during a denial of service attack.

6.2 Observation window

The observation window corresponds to the number of packets (window size L) that will be analyzed by the mechanism. Within each observation window w_i, the frequency at which the field to be observed appears is computed in $O(L^2)$ time. The probability of occurrence of the fields of interest is measured as follows:

$$p_j = \frac{o_j}{L}, \tag{3}$$

where $o_j, (1 \leq j \leq L)$ is defined as the frequency of the observed fields in w_i, $|w_i| = L$ is the number of packets in w_i and p_j is the probability of occurrence of the field observed in w_i.

6.3 Statistical method

The statistical method consists in applying one of the methods defined by the mechanism, chi-square (Equation 1) or entropy (Equation 2), defined as δ, from the observed frequency $(o_1 \dots o_j)$ and the probability of occurrence of the field $(p_1 \dots p_j)$ in the window (w_i). The calculated values are stored in a list containing the previously calculated values. These values will serve as the basis for setting the detection threshold.

6.4 Detection threshold

The detection threshold T_{w_i} defines the sensitivity of the proposed mechanism. The calculation of the threshold is given as:

$$T_{w_i} = \begin{cases} T_{w_i}, & i = 1 \\ T_{w_i}\alpha + (1-\alpha)T_{w_{i-1}}, & i > 1 \end{cases} \tag{4}$$

where T_{w_i} represents the current value of the applied anomaly method calculated, T_{w-1} represents the estimated value in the previous window (w_{i-1}) without attack and α is the smoothing coefficient. The coefficient α assumes values $(0 \leq \alpha \leq 1)$, then that the closer to 1, the greater the weight applied to recent information and the

closer to 0 the greater the weight applied to past information. In this way, the threshold can be adjusted according to the characteristics of the network to reduce the incidence of false alarms.

6.5 DoS Attack Detection Mechanism

Based on the components above, this work presents a new statistical DoS attack detection mechanism for an SDN architecture. Note that, unlike other works presented in the literature, the proposed strategy can adjust the DoS detection threshold limits at runtime. The proposal is presented in Algorithm 1.

Algorithm 1 *DoS Attack Detection Mechanism*

Input: T_{w_1}, L, α

1: $i \leftarrow 2$
2: $attack \leftarrow$ "$false$"
3: **while** *true* **do**
4: $w_i \leftarrow \{\}$
5: **while** $|w_i| \leq L$ **do**
6: $f \leftarrow$ extract the observed field from *packet_in*
7: $w_i \leftarrow w_i \cup \{f\}$
8: **end while**
9: Compute o_j and $p_j \; \forall j \in w_i$
10: $\delta \leftarrow StatisticalMethodResult(w_i)$
11: **if** $!attack$ **then**
12: $T_{w_i} \leftarrow T_{w_i}\alpha + (1-\alpha)T_{w_{i-1}}$
13: **end if**
14: **if** $\delta > T_{w_i}$ **then**
15: $attack \leftarrow$ "$true$"
16: **else**
17: $attack \leftarrow$ "$false$"
18: $i \leftarrow i + 1$
19: **end if**
20: **end while**

Initially, the controller receives as input the initial threshold value (T_{w_1}), window size (L) and smoothing coefficient (α). It is assumed that in the first window there is no attack and that the first threshold (T_{w_1}) has already been calculated (lines 1 – 2). The controller starts a loop that always keeps analyzing all *packet_in*'s (line 3). As packets enter the OpenFlow controller input queue, the controller start window as empty (line 4) and then extracts the field to be parsed by the proposed detection engine (lines 5 – 8). The parsed field corresponds to a predefined field from the available header fields in the packets. The observation window w_i is composed of these fields until they reach a predetermined size.

In this article, we employ sizes of 20, 60, and 100 packets parsed for the source IP address field, as we will see in Section 7. When the observation window is filled, that is, when $|w_i| = L$, the observed frequency ($o_1 \ldots o_j$) and the probability of occurrence ($p_1 \ldots p_j$), for all fields in the window (w_i), as calculated in (line 9). With all observed frequency values and all calculated probabilities, it is possible to perform the statistical analysis defined for the detection mechanism (chi-square or entropy) in w_i, where δ is defined (line 10). If the variable denoting the attack is set to $false$ (line 11), then the window threshold (T_{w_i}) is updated. If the value of the statistical method is above the threshold ($\delta > T_{w_i}$)[1] the DoS attack is detected (lines 14 – 16), otherwise, it will be considered as legitimate traffic ($attack =$ "$false$"), increments i and the loop runs again.

6.6 Complexity Analysis

The time complexity of the proposed solution is analyzed as follows. To calculate the frequency of the observed field, it is necessary to analyze the entire observation window, of size m, for each element, so its time complexity is $O(m^2)$. To calculate chi-square or entropy it is necessary to calculate the number of distinct occurrences in the window. So in the worst case, we have window size m, thus $O(m)$. Thus, overall, the total time complexity of window detection analysis is $O(m^2)$.

7 Experiments

In order to evaluate the proposed architecture, we implemented the proposed detection model and analyzed its performance. The following sections present the experimental environment, the methodology, the metrics considered in our analysis and the results observed in the experiments.

7.1 Experimental Setup and Methodology

The experiments were performed using Mininet as an emulator of an SDN architecture [26]. In Mininet the switches and controllers operate under the OpenFlow protocol enabling the implementation of SDN applications. The controller used was POX, which enables rapid prototyping in the Python language to create features in the network [27]. The controller has been implemented to collect source IP addresses of new incoming packets. A function was developed for calculating packet statistics. The experiments were performed on the Ubuntu 16.04 LTS operating system, using a core i5 7400 3.0 GHz CPU with 4Gb of RAM.

The proposed detection mechanism is not influenced by the size of the network, but by the volume of packets that make up the observation window. To this end, using a network topology was considered in order to allow balancing the workload between the victim and the attacker. The topology consists of 4 OpenFlow switches, 9 hosts and 1 controller POX. The IPs range from 10.0.0.1 to 10.0.0.9, with host 10.0.0.3 being the victim server and host 10.0.0.7 the attacker. The range of spurious IPs are generated at random. The bandwidth of each link is 100 Mbps, as shown in Figure 2. The victim's host is located on the "Switch OpenFlow 2" network and the DoS attacker performs the attack from the "Switch OpenFlow 4" network.

The generation of legitimate packets is performed using the Scapy [28]. Scapy is a tool that allows creating TCP and UDP packets. The spoofed SYN packets are generated using the tool Hping [29]. This tool is used for security testing on networks operating as a packet generator for the TCP/IP protocol. Two types of

[1]The relationship between the δ and T_{w_i} depends on the field to be analyzed. For example, if the field is the source IP, use $\delta > T_{w_i}$, if the field is the destination IP, use ($\delta < T_{w_i}$).

traffic were generated: normal traffic and spurious traffic. Normal traffic is used to obtain the statistical values used by the detection mechanisms of anomalies (chi-square and entropy) under normal conditions. With this, it is possible to define the detection threshold, based on Equation 4. Spurious traffic is injected to verify the behavior of the solution.

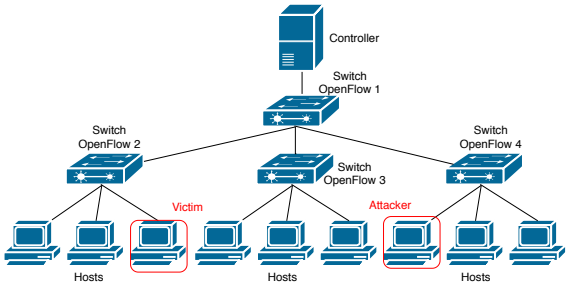


Figure 2: Network topology.

To analyze the proposed solution, SYN Flood attack were simulated by calculating the entropy and chi-square values using source IP addresses with different observation window sizes and smoothing coefficients (threshold) to identify the attack traffic. In the experiments, 20,500 packets were considered. The observation windows had the following sizes: 20, 60 and 100 packets. The window is moved sequentially with the same number of packets from the previous window without overlap, as shown in Figure 3. The smoothing coefficients (α) evaluated were: 0.1, 0.5 and 0.9. The simulation time of each experiment was 200 s, and in the hundredth second the SYN Flood attack was started at a rate of 250 packets per second from a host belonging to the "Switch OpenFlow 2" network. The rate of legitimate packets is 5 packets per second. Five simulations of each method were performed to analyze the behavior of the solution.

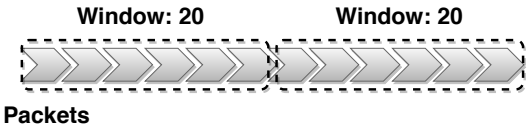


Figure 3: Moving windows

Figures 4 and 5 show the outputs of the chi-square and entropy values. Such values are obtained by Equations (1) and (2), respectively. The observed field was the source IP address, the observation windows have the size of 60 packets and the smoothing coefficient of 0.5 for the detection threshold.

The observation period shows 200 seconds of analysis, which at a rate of 250 packets per second generates about ≈833 windows of 60 packets. Before the onset of the attack, the chi-square and entropy measurements are in the range of 180 to 220 for the chi-square and 0.44 to 0.49 for the entropy for the normal traffic. During the attack, there is a considerable increase in these values in both techniques. The reason for this is that the source IP address is changed in each new packet, due to the attack characteristic, which causes the source IP address quantities to increase substantially filling the entire observation window, causing an anomaly. In the case of the chi-square, a peak is observed at the beginning of the attack injection, but during and after the attack, the values return to a normal traffic situation. entropy, in turn, best classifies the beginning and end of the attack, differentiating the attack period from a normal traffic situation. It is important to note that the chi-square is asymmetric, generating a peak only at the beginning of the attack, after the end of the attack there is a slight change between the observed values.

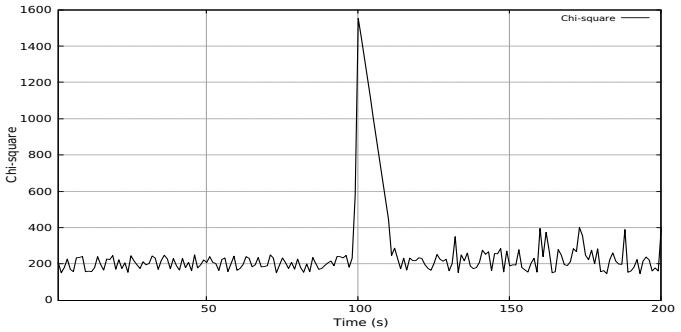


Figure 4: DoS attack identification using chi-square.

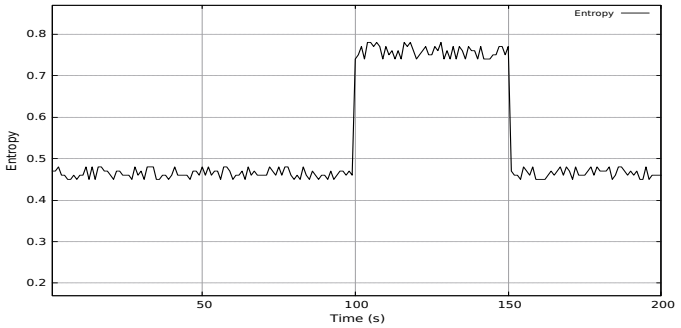


Figure 5: DoS attack identification using entropy.

It can be observed that both methods can identify anomalies in the network based on the deviation of the traffic distribution. The two methods were compared to identify which method has the highest detection rate and the lowest false alarm rate in the proposed SDN architecture model.

7.2 Metrics

The quality of the results was measured using a confusion matrix. Table 1 exemplifies the matrix model used. The matrix is composed of predicted values (results obtained) and actual values (expected results) of normal traffic and spurious traffic, based on the rates of false negative (FN), false positive (FP), true negative (TN) and true positive (TP).

Table 1: Confusion matrix.

Predicted	**Actual**	
	Attack	**Normal**
Attack	TP	FP
Normal	FN	TN

The FN rate is given by the number of packets classified as normal (negative) but that are confirmed as attack packets (positive). The FP rate is given by the number of packets classified as attack packets (positive) but they are confirmed as normal (negative) packets. On the other hand, TN rate indicates the number of packets correctly classified as legitimate (negative) and TP rate represents the number of packets correctly classified as attack packets (positive).

In this work, we use precision, recall and $F1$-score to evaluate the effectiveness of the proposed DoS detection mechanisms.

Precision represents the proportion of positive identifications that were correctly identified. This metric measures the quality of the results and is defined as follows:

$$\text{Precision} = \frac{TP}{TP + FP}. \tag{5}$$

Recall is the proportion of positives that have been identified correctly, measuring the completeness of the mechanism. This metric is defined as follows:

$$\text{Recall} = \frac{TP}{TP + FN}. \tag{6}$$

F1-score is defined as a harmonic mean between precision and recall, aiming to bring the balance between both, to indicate the overall performance of the technique. High F1-score values indicate that the detection technique presents more accuracy. this metric is defined as follows:

$$\text{F1} = \frac{(2 * TP)}{(2 * TP + FP + FN)}. \tag{7}$$

7.3 Results

The confusion matrix for the entropy and the chi-square can be seen in Table 2 and 3, respectively. The data in the table was obtained based on the simulation results for each statistical method presented. The smoothing coefficient can be observed immediately to the left at the beginning of each method analyzed.

The smoothing coefficient plays an important role in the definition of detection thresholds since it is responsible for defining the detection sensitivity of the proposed method. Through it is possible to assign weights to the values and thus the threshold represents an exponential weighted moving average, whose latest values have more weight. A very low coefficient may leave the solution more sensitive to small variations in the values obtained, providing a higher detection, but this sensitivity may cause an increase in the FP rate. On the other hand, a very high coefficient provides greater freedom for the change in the values obtained and consequently a lower detection resulting and an increase in the rate of FN. It is therefore ideal to seek a balance for the smoothing coefficient in order to provide it with an appropriate threshold that reflects well the characteristics of the network and can reduce these rates while maintaining an acceptable level of accuracy.

The size of the window is related to the accuracy of attack detection. We can note that the larger the window, the greater the accuracy in detecting the attack, this is due to the large number of packets that fill the window reflect in a greater representativity of the data to be observed, however this size can cause a greater consumption of resources by part of the controller and slower detection because the number of packets the solution needs to fill the window is larger. For a small window, we can observe a low precision in the results, because the observed values represent a small part of the total set to be analyzed, in contrast, we obtain a faster detection, providing a lower consumption of resources.

Considering the analyzed parameters of the smoothing coefficients and observation windows, the mechanisms had similar behaviors, although the entropy was 0.59% more accurate on average than the chi-square, due to the reduction of the FP and FN rate in both cases.

Figure 6 shows the result of the Precision. We can see that the mechanism can correctly identify spurious traffic in an acceptable way, avoiding false positives. Figure 7 shows the result of the Recall. The results show that the mechanism can effectively separate what is spurious traffic from legitimate traffic.

Figure 8 shows the result of the F1 score. Although entropy provides higher accuracy when compared to chi-square results, it requires a larger packet volume to detect any change in traffic, which can increase detection time. For example, for the 60-packet window with a smoothing coefficient of 0.1 ($\alpha = 0.1$), the entropy needed to analyze approximately 582 packets to detect the attack, the equivalent of ≈10 windows, while the chi-square needed only 478 packets, the equivalent of ≈8 windows, reducing the volume of packets evaluated by about 17%, as shown in Figure 9. This is because chi-square allows the comparison between two consecutive distributions, allowing to detect much changes faster. Unlike entropy, which represents the dispersion of several values in a given probability distribution. In summary, chi-square has reduced on average by 14.89% of the number of packets to be analyzed in the observation windows used.

It is important to note that the results presented do not take into account the attack rate, because the proposed solution focuses on analyzing the behavior of the observation windows and the sensitivity of the detection thresholds, finding the ideal balance between these components in order to meet the needs of the network operator.

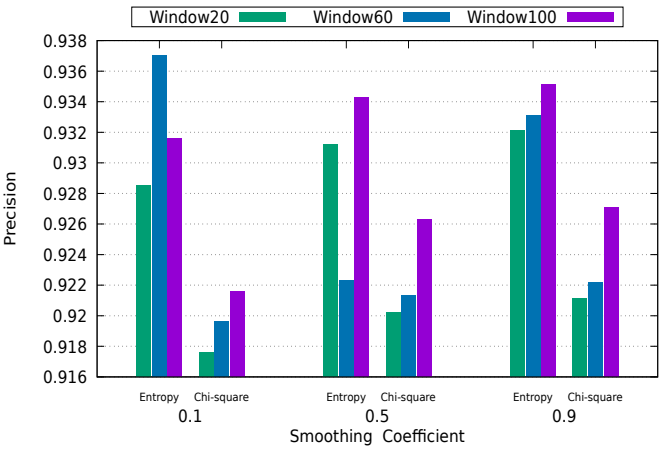


Figure 6: Precision results.

Table 2: Entropy confusion matrix.

$\alpha = 0.1$

Window 20		
Predicted	**Actual**	
	Attack	**Normal**
Attack	95.5%	8.9%
Normal	4.5%	91.1%

Window 60		
Predicted	**Actual**	
	Attack	**Normal**
Attack	95.7%	8.6%
Normal	4.3%	91.4%

Window 100		
Predicted	**Actual**	
	Attack	**Normal**
Attack	95.9%	8%
Normal	4.1%	92%

$\alpha = 0.5$

Window 20		
Predicted	**Actual**	
	Attack	**Normal**
Attack	95.1%	7.2%
Normal	4.9%	92.8%

Window 60		
Predicted	**Actual**	
	Attack	**Normal**
Attack	95.5%	6.7%
Normal	4.5%	93.3%

Window 100		
Predicted	**Actual**	
	Attack	**Normal**
Attack	95.6%	6.6%
Normal	4.4%	93.4%

$\alpha = 0.9$

Window 20		
Predicted	**Actual**	
	Attack	**Normal**
Attack	94.4%	6.6%
Normal	5.6%	93.4%

Window 60		
Predicted	**Actual**	
	Attack	**Normal**
Attack	94.7%	6%
Normal	5.3%	94%

Window 100		
Predicted	**Actual**	
	Attack	**Normal**
Attack	94.8%	5.9%
Normal	5.2%	94.1%

Table 3: Chi-square confusion Matrix

$\alpha = 0.1$

Window 20		
Predicted	**Actual**	
	Attack	**Normal**
Attack	95%	9.7%
Normal	5%	90.3%

Window 60		
Predicted	**Actual**	
	Attack	**Normal**
Attack	95.2%	9.6%
Normal	4.8%	90.4%

Window 100		
Predicted	**Actual**	
	Attack	**Normal**
Attack	95.6%	9.3%
Normal	4.4%	90.7%

$\alpha = 0.5$

Window 20		
Predicted	**Actual**	
	Attack	**Normal**
Attack	94.7%	7.6%
Normal	5.3%	92.4%

Window 60		
Predicted	**Actual**	
	Attack	**Normal**
Attack	95.1%	7.1%
Normal	4.9%	92.9%

Window 100		
Predicted	**Actual**	
	Attack	**Normal**
Attack	95.3%	7%
Normal	4.7%	93%

$\alpha = 0.9$

Window 20		
Predicted	**Actual**	
	Attack	**Normal**
Attack	93.8%	6.9%
Normal	6.2%	93.1%

Window 60		
Predicted	**Actual**	
	Attack	**Normal**
Attack	94%	6.5%
Normal	6%	93.5%

Window 100		
Predicted	**Actual**	
	Attack	**Normal**
Attack	94.1%	6.3%
Normal	5.9%	93.7%

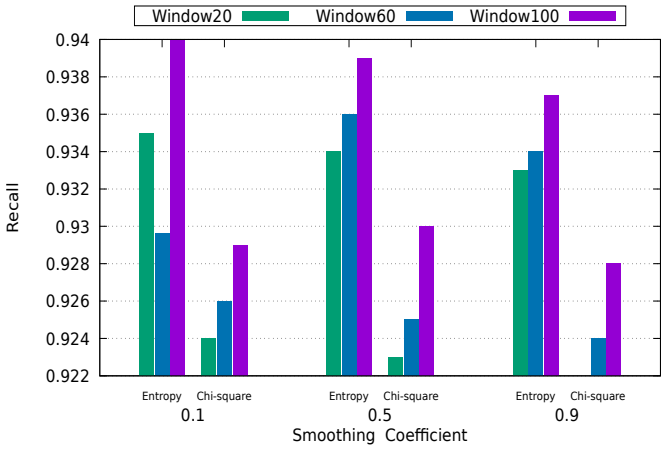


Figure 7: Recall results.

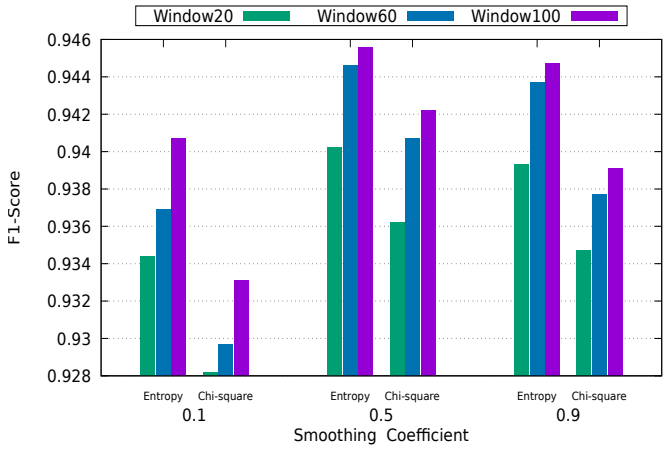


Figure 8: F1 Score results.

Finally, Figure 10 shows the result of time detection. The smoothing coefficient plays an important role in this regard as it allows manipulating the sensitivity of the mechanism which can result in faster or slower detection. For example, coefficient 0.1 provides more immediate alerts, unlike coefficient 0.9, at the cost of higher FP and FN rates. This is due to the weight of information given to the mechanism for calculating the detection threshold.

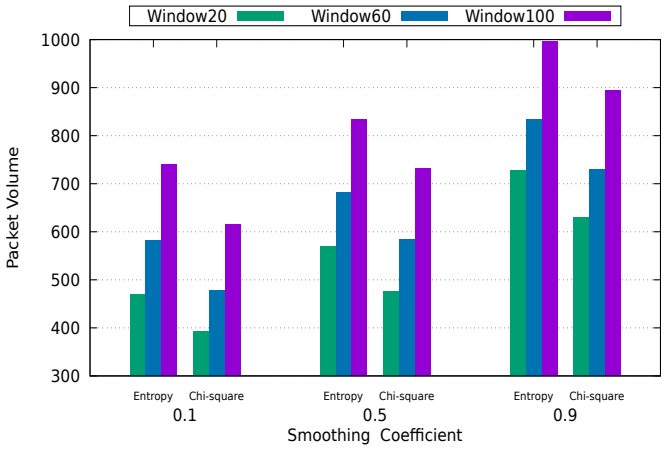


Figure 9: Packet volume.

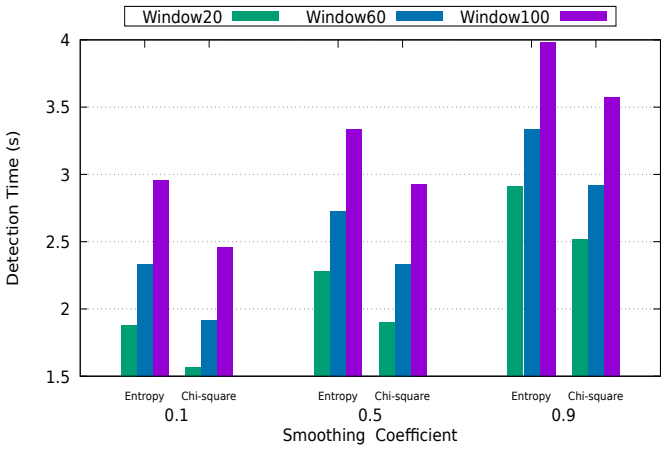


Figure 10: Detection Time.

8 Conclusion

This paper proposes a solution for the detection of DoS attacks based on the use of any mechanism for the detection of network anomalies. The mechanism works alongside the controller. Two techniques are evaluated for the detection of DoS attacks: entropy and chi-square. Both techniques were evaluated in terms of the rate of false positives and false negatives. From the experimental results, chi-square and entropy could detect the TCP-SYN Flood attack with an acceptable accuracy rate. The entropy was a little more accurate due to the reduction in the number of false positives and false negatives, around 0.59%, while the chi-square detects the anomaly with a smaller number of packets, on average 14.89% less entropy. One of the main benefits of the proposed solution is its versatility. Any parameter in the solution can be changed to suit the characteristics of the network. The detection field (i.e, source IP, destination IP, source port and source port), detection threshold and window size can be configured to meet SDN requirements. As future work, we want to handle this information to establish mitigation strategies in SDN architectures under different scenarios and different attack rates.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment This work is partially supported by the MCTIC/RNP/CTIC (Brazil) through the project P4Sec.

References

[1] D. Kreutz, F. Ramos, P. Verssimo, and C. Rothenberg, "Software-Defined Networking: A Comprehensive Survey," *IEEE*, 2014.

[2] D. Kreutz, F. M. Ramos, and P. Verissimo, "Towards Secure and Dependable Software-defined Networks," in *Proceedings of the Second ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking*, ser. HotSDN '13. New York, NY, USA: ACM, 2013, pp. 55–60. [Online]. Available: http://doi.acm.org/10.1145/2491185.2491199

[3] A. Lara, A. Kolasani, and B. Ramamurthy, "Network innovation using openflow: A survey," *IEEE communications surveys & tutorials*, vol. 16, no. 1, pp. 493–512, 2014.

[4] D. M. F. M. e Otto Carlos Muniz Bandeira Duarte, "Authflow: authentication and access control mechanism for software defined networking," in *Ann. Telecommun*, vol. 71, 12 2016, pp. 607 – 615.

[5] F. Lau, S. H. Rubin, M. Smith, and L. Trajkovic, "Distributed denial of service attacks," vol. 3, 02 2000, pp. 2275 – 2280 vol.3.

[6] S. Tritilanunt, S. Sivakorn, C. Juengjincharoen, and A. Siripornpisan, "Entropy-based input-output traffic mode detection scheme for DoS/DDoS attacks," in *2010 10th International Symposium on Communications and Information Technologies*, Oct 2010, pp. 804–809.

[7] A. Bhandari and A. Sangal, "Destination Address Entropy based Detection and Traceback Approach against Distributed Denial of Service Attacks," *I. J. Computer Network and Information Security*, vol. 8, pp. 9–20, jul 2015.

[8] K. Kurihara and K. Katagishi, "A Simple Detection Method for DoS Attacks Based on IP Packets Entropy Values," in *2014 Ninth Asia Joint Conference on Information Security*, Sep. 2014, pp. 44–51.

[9] J. Li, Y. Liu, and L. Gu, "DDoS attack detection based on Neural Network," in *2010 2nd International Symposium on Aware Computing*, Nov 2010, pp. 196–199.

[10] C. Shannon, "A mathematical theory of communication," vol. 27, pp. 379–423, 1948.

[11] M. Broniatowski and S. Leorato, "An estimation method for the Neyman chi-square divergence with application to test of hypotheses," *Journal of Multivariate Analysis*, vol. 97, no. 6, pp. 1409 – 1436, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0047259X06000194

[12] A. Taha, "Software-Defined Networking and its Security," 2014.

[13] M. Casado, T. Koponen, R. Ramanathan, and S. Shenker, "Virtualizing the Network Forwarding Plane," *Workshop on Programmable Routers for Extensible Services of Tomorrow*, 2010.

[14] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "OpenFlow: Enabling Innovation in Campus Networks," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 2, pp. 69–74, Mar. 2008.

[15] O. N. Foundation, "Openflow switch specification 1.0," vol. 1, pp. 1–44, december 2009.

[16] R. A. J. Gondim, "Mirror Saturation in Amplified Reflection DDoS," *JNIC-Jornadas Nacionales de Investigacin en Ciberseguridad*, 2019.

[17] R. Bani-Hani and Z. Al-Ali, "Syn flooding attacks and countermeasures: A survey," 04 2013.

[18] H. Beitollahi and G. Deconinck, "Analyzing well-known countermeasures against distributed denial of service attacks," *Computer Communications*, vol. 35, no. 11, pp. 1312 – 1332, 2012.

[19] J. Tajer, A. Makke, O. Salem, and A. Mehaoua, "A comparison between divergence measures for network anomaly detection," in *7th International Conference on Network and Service Management*, 2011.

[20] L. Feinstein, D. Schnackenberg, R. Balupari, and D. Kindred, "Statistical approaches to DDoS attack detection and response," in *Proceedings DARPA Information Survivability Conference and Exposition*, vol. 1, April 2003, pp. 303–314 vol.1.

[21] S. Sharma, S. K. Sahu, and S. K. Jena, "On selection of attributes for entropy based detection of DDoS," in *International Conference on Advances in Computing, Communications and Informatics*, 2015.

[22] F. Leu and C. Pai, "Detecting DoS and DDoS Attacks Using Chi-Square," in *2009 Fifth International Conference on Information Assurance and Security*, vol. 2, Aug 2009, pp. 255–258.

[23] S. Oshima, A. Hirakawa, T. Nakashima, and T. Sueyoshi, "DoS/DDoS Detection Scheme Using Statistical Method Based on the Destination Port Number," in *2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Sep. 2009, pp. 206–209.

[24] A. C. Lapolli, J. A. Marques, and L. P. Gaspary, "Offloading Real-time DDoS Attack Detection to Programmable Data Planes," in *IFIP/IEEE International Symposium on Integrated Network Management*, 2019.

[25] R. Carvalho, J. L. Bordim, and E. A. P. Alchieri, "Entropy-based DoS Attack identification in SDN," *21st Workshop on Advances in Parallel and Distributed Computational Models*, 2019.

[26] M. Team, "Mininet - an instant virtual network on your laptop (or other pc)." [Online]. Available: http://mininet.org/

[27] M. MC, "NOX." [Online]. Available: https://github.com/noxrepo/pox

[28] Scapy, "Packet crafting for Python2 and Python3." [Online]. Available: https://scapy.net/

[29] S. Sanfilippo, "Hping - Active Network Security." [Online]. Available: http://www.hping.org/

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 216-219 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Development of Evaluation Metrics for Learners in Unplugged Activity

Woochun Jun*

Seoul National University of Education, Department of Computer Education, 06639, Korea

ARTICLE INFO

Article history:
Received: 14 July, 2019
Accepted: 04 November, 2019
Online: 05 December, 2019

Keywords:
Unplugged Activity
Unplugged Computing
Evaluation Metrics

ABSTRACT

As well as physical computing, unplugged activity is important in ICT (Information and Communication Technology) education, that uses information and communications technology to support, enhance, and optimize the delivery of information, and software education. In order to achieve educational objectives in unplugged activity, evaluation metrics of learners are necessary. However, there is little work on evaluation metrics in unplugged activity. In this paper, evaluation metrics of learners are proposed and developed using based on the general design principles and opinions of experts.

1. Introduction

This paper is an extension of work originally presented in ICTC2018[1].

Unplugged activity is a teaching method that literally learns about a computer without using a computer. In the origin, Professor Tim Bell of New Zealand suggested a new way of teaching students to learn the principles of computer science by learning the principles of computer operation and creative ideas that invented and developed computers through specific manipulation activities. Through the unplugged method, play learning materials in the form of specific manipulations were developed and applied to understand the principles of operation of algorithms and computers, such as expression, alignment, search, algorithm, routing, and deadlock of data without using the computer.

Unplugged learning is a new method of education for elementary and middle school computer education, so students can learn the principles of computer science easily and interestingly. Unplugged learning was initially a teaching method for elementary school students, but it can be easily used and applied to anyone who wants to learn the principles of computer science. Learning in the form of play related to specific experiences and real life has led to a change in the concept of computer science as an easier and more interesting subject. And it gave teachers a new direction in computer education to effectively teach logical computer science subjects without the need for computer and programming.

Unplugged activity has a big advantage. That is, it is to let students not to feel busy, feel fun and good. It also emphasizes cooperative thinking rather than competition from individuals, and uses computers to help people think in abundance. Unplugged activity can be said to be a new computer education created between ideas and activities for learners, and it has given teachers a new direction in computer education in that they can effectively teach logical computer science subjects without the need for computer and programming learning [2,3].

For the successful fulfillment of unplugged activity in ICT education, it is necessary to establish certain and secure evaluation standards or metrics for learners. That is, according to proper evaluation metrics, learners must be evaluated and education contents or teaching-learning method can be updated or modified depending on result of evaluation metrics.

However, in the literature, there is little research work on evaluation of learners in unplugged activity. The purpose of this paper is to develop evaluation metrics of learners in unplugged activity. For this purpose, in this work, overall categories are proposed and detailed evaluation metrics items are proposed depending on each category.

The rest of this paper is organized as follows. In Chapter 2, related works for unplugged activity is introduced. In Chapter3, evaluation metrics of learners are proposed. Finally, in Chapter 4, conclusions and further research issues are discussed.

*Corresponding Author: Woochun Jun, Seocho Dong, Seocho Gu, Seoul, Korea, wocjun@snue.ac.kr

2. Related Works

2.1. Principles of Unplugged Activity

In the literature [4], there are 10 principles of unplugged activity. That is, 10 class activity design principles are suggested. Those principles are also main characteristics of unplugged activity. The followings are those principles.

- No Computers Required

The first principle implies that a computer is not used at all. It means that the main objective of unplugged activity is not to use computers while students are learning the principles and concepts of computer science.

- Real Computer Science

Unplugged activity can have students to know basic principles of computer science such as database, software engineering, computer architecture, computer graphics, communication network, operating systems, and programming languages through a variety of unplugged activities. That is, the purpose of unplugged activity is to let students understand the various concepts of computer science. It means that unplugged activity must include the main principles of computer science in their activities.

- Learning by doing

"Learning by doing" is accompanied by the emphasis of progressive education. It provides solutions based on the constructivism that students can learn new concepts by participating in active activities voluntarily with a purpose. That is, students can learn the concepts of computer science through discovery learning on their own. That is, rather than providing what learners need to know from the beginning, such as the discovery learning principle presented by Brunner, the learning can proceed in a situation that requires the learner to organize the final form or content on his or her own. In other words, the logic is that students can understand their learning more deeply by participating in the discovery process of knowledge. Therefore, it does not conform to the basic concepts of unplugged activities to present and understand concepts, principles and ideas immediately.

- Fun

Unplugged activity should keep the rules and principles by learning the concepts of computer science through play and activity, gain fun through play, and make learners feel a sense of achievement in your activities. This is because if students are not given enough fun in the course of "Learning by doing," they can turn into activities that the teacher understands and forces them to understand rather than acquire them. Therefore, teachers present some rules along with basic data, and students solve tasks according to the rules presented by the teachers. The caveat in unplugged activities should be that the activities give students challenges and competitive psychology, as well as fully reflecting the interesting points in the process of problem solving. The activities should also be organized so that students can be fully satisfied with the sense of achievement they achieve in the process.

- No special equipment

The materials available for use in unplugged activities can be various objects used in our lives. That is, we do not use computers, and we use things that are readily available around us to organize activities that include the concepts of computer science. For instance, various materials can be used, such as cards, threads, chalk, cups, balls, pencils, and colored paper.

- Variation encouraged

Unplugged activity recommends that computer scientists from all over the world participate together through the Creative Commons License (CCL). It also respects open thinking so that changes, adaptations and extensions can be made to previously developed activities.

- For Everyone

Unplugged activity can be applied to any age group. If you are a beginner who first comes across the principles of computer science, you can use them extensively, from elementary school students to the elderly, regardless of age. It can also be used in various ways, not dependent on a particular country, culture or society.

- Co-operative

One of unplugged activity's main goals is cooperation. In other words, rather than solving a problem on the learner's own, he or she solves the problem through various collaborative work with his colleagues, and discovers the principles of computer science contained in it. That is, it emphasizes collaborative learning, which is the main philosophy of the theory of constructivism of learning. It is suitable for team activities rather than individuals, and is applicable in competitive situations.

- Stand-alone activities

The main objective of unplugged activities is to acquire the concepts of computer science embedded in the content of the activities, while carrying out activities that include the basic concepts or principles of computer science. Therefore, individual activity should not be dependent each other and can present a single computer principle in one activity. In other words, if multiple concepts or principles are included in an activity, it is difficult to exclude the possibility of a misconception.

- Resilient

Since unplugged activity is focused on activities that can learn the concepts or principles of computer science and also does not focus on understanding or memorization, it is difficult to present the principles or all very difficult steps precisely. Therefore, even if there is the possibility of small mistakes or misconceptions, it should be able to demonstrate the flexibility to implement them, unless they are a major obstacle to understanding the principles of computer science.

2.2. Literature Review

In [5], unplugged activities for information protection education applicable to primary and secondary schools are proposed. Unplugged activities developed in this study were designed using Lickona's integrated morality model and unplugged activity design pattern, and developed after expert review and pilot testing for elementary school students. The developed activities were applied to 21 middle school students and their usefulness was assessed.

In [1], an evaluation standard of unplugged activity is proposed. The following table 1 shows the whole standard.

Table 1: Evaluation Standard

Number	Standards	Description
1	Computer	Is any computer (including notebook, tablet PC, or other electronic devices) necessary?
2	Fun	Does the activity give fun to students?
3	Computer Science Principle	Does the activity deliver any idea or principle of computer science theory?
4	Student Participation	Does the activity require active participation of all students?
5	Independent Activity	Does the activity represent any independent computer science theory?
6	Learning by Doing	Does student learn by doing?
7	For everyone	Can everyone participate the activity?
8	Background	Does the activity require any academic background or preliminary course?
9	Safety	Is the activity is safe?
10	Report	Does the activity let student submit activity report?

3. Development of Evaluation Metrics for Learners

In this chapter, evaluation metrics for learners in unplugged activity are proposed. The proposed metric is developed based on the previous research works and experts from ICT education such as professors and teachers.

3.1 Design Principles

The proposed evaluation metrics are developed as following principles. First, they are based on 10 design principles of unplugged activity in previous chapter. It means that the metrics are in accordance with 10 design principles. Second, the proposed metrics are developed to test students' participation and interest in unplugged activity. Third, metrics are developed to test how much learners understand the overall idea or principle in computer science theory.

3.2 Evaluation Metrics

The following table 2 shows the overall evaluation metrics of unplugged activity.

Table 2: Evaluation Metrics

Number	Metrics	Description
1	Understanding	Do learners understand the whole idea or principle of computer science theory?
2	Fun	Do learners have fun or feel fun during unplugged activity?
3	Participation	Do learners participate in activity actively or positively?
4	Interest	Do learners have interests on unplugged activity?
5	Safety	Do learners feel safe during unplugged activity?
6	Cooperation	Are learners engaged well in unplugged activity?
7	Assistance	Do learners have chance for assistance from teacher or friends?
8	Communication	Do learners have chance to communicate each other during unplugged activity?
9	Expandability	Can learners expand the unplugged activity with more member?
10	Applicability	Can learners apply the unplugged activity for other similar activity?

4 Conclusions and Further Research Issues

Currently informatics education becomes popular in the forms of software education and ICT education, etc. Informatics education is important in raising human resource in artificial intelligence areas [6,7]. Informatics education also becomes mandatory education in primary and secondary schools in many countries.

Unplugged activity becomes popular in ICT education and software education since it does not require any special equipment including computer. It means that unplugged activity can be done any place any time. The main purpose of unplugged activity is to let learners' study computer science theory or

principles with play activity.

In order for unplugged activity to be successful, the correct and general evaluation metrics for learners are necessary. It is very important to check how learners understand the overall principle or theory. However, in the literature, there is little research work on evaluation metrics for learners. In this paper, evaluation metrics are developed. The metrics are developed in accordance with design principles of unplugged activity and the general education purpose of ICT education.

The further research issues are as follows. First, it is necessary to refine the proposed evaluation metrics further. That is, more detailed metrics are necessary depending on various unplugged activities. Second, it is also necessary to test the validity of the proposed metrics. For this purpose, the extensive survey work and analysis work are essential.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] W. C. Jun," A Study on Development of Evaluation Standards for Unplugged Activity" in 2018 International Conference on Information and Communication Technology Convergence, Jeju Island, Korea, 2018. http://10.1109/ICTC.2018.8539505

[2] J. R. Kim, "A Study on Systematic Review of Unplugged Activity" Journal of the Korean Society of Information Education, 22(1), 103-111, 2018. http://dx.doi.org/10.14352/jkaie.2018.22.1.103

[3] Y. H. Bae, C. S. Na, "Effectiveness of Unplugged Activity in Korean Education: A Meta-Analysis" The Journal of Educational Information and Media, 25(1), 121-150, 2019. http://dx.doi.org/10.15833/KAFEIAM.25.1.121

[4] W. K. Lee and J. M. Kim, "Research Report on Unplugged Activity", Research Report, The Korean Society of Computer Education, 2015.

[5] Y. J, Jang, D. H. Kim, H. S. Kim, W. K. Lee, H. C. Kim, "Development of Unplugged Activity and its Evaluation of Usability for Information Security Education" The Journal of Korean Association of Computer Education, 14(1), 55-67, 2011.

[6] C. Gokul, P. Sakthivel and P. R. Karthikeyan, Minimization of Test Time in System-on-Chip using Artificial Intelligence based Test Scheduling Techniques, Neural Computing & Applications, 2019. https://doi.org/10.1007/s00521-019-04039-6.

[7] C. Gokul, P. Sakthivel and P. R. Karthikeyan, Test scheduling for system on chip using modified firefly and modified ABC algorithms, SN Applied Sciences. 1(9), 1079, 2019. https://doi.org/10.1007/s42452-019-1116-x.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 220-228 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

SCMS: Tool for Assessing a Novel Taxonomy of Complexity Metrics for any Java Project at the Class and Method Levels based on Statement Level Metrics

Issar Arab[1,2,*], Bouchaib Falah[3], Kenneth Magel[4]

[1]*Department of Informatics, Technical University of Munich, 80333, Germany*

[2]*Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta, 8613, Canada*

[3]*School of Science and Engineering, Al Akhawayn University in Ifrane, 53000, Morocco*

[4]*Faculty of Computer Science, North Dakota State University, ND 58108-6050, USA*

ARTICLE INFO

Article history:
Received: 20 September, 2019
Accepted: 14 November, 2019
Online: 05 December, 2019

Keywords:
Software Testing
Complexity Metrics
Software Metrics Tool
Effectiveness
Data Flow
Data Usage
Taxonomy
JAVA

ABSTRACT

Software is the primary and indispensable entity in our technologically driven world. Therefore, quality assurance, and in particular software testing, represents a vital component in the software development cycle. Throughout the years, many tools have been developed to collect metrics of software that had been implemented. These tools have some differences but continue to play an important role in improving the quality of software products. This paper introduces a newly developed tool, named Spectra Complexity Metrics System (SCMS), which compiles a novel taxonomy of complexity metrics of any given software written in the Java programming language. Our suggested metrics have been invented to identify and evaluate the characteristics of Java computer programs. They aim at increasing the efficiency of the testing process by significantly reducing the number of test cases without having a significant drop in test effectiveness. We assess our proposed taxonomy of different complexity metrics based on the product levels (statement, method, and class) and their characteristics. For further evaluation, our software metrics coverage is compared to other existing software metric tools. The results show the novelty of our taxonomy of complexity metrics and the capability of our tool to compute these measurements based on all three of the product level categories. We have published our tool at https://github.com/issararab/SCMS under an open-source license.

1. Introduction

Over the current decade, concerns over present and future software quality have grown, as have the range and complexity of software and its applications. Increasingly, developers, researchers, and users are dissatisfied with the quality of available software. Hence, we have seen a growing focus on software testing, where the engineers' main job is to assess and quantify the quality of a given product. The key challenge in this field is to reduce costs and maximize benefits. Those challenges have motivated software engineers to develop metrics or rules for quantifying given characteristics and attributes of software entities. For good quality software, these characteristics should be understandable and measurable [1]. These metrics aim to validate and verify the quality and the assurance of software. As it is extremely time-consuming and requires a great deal of time and effort, the overall testing process needs to be automated for testing practices [2].

This article is an extension of a paper previously presented in the International Conference on Software Engineering Research and Practice (SERP 2015), in Las Vegas, NV, USA [3]. Our contribution in this paper consists of explaining the theory behind our suggested taxonomy and presenting the first tool, SCMS, that compiles this introduced set of metrics. Furthermore, we are benchmarking the software metrics coverage against state-of-the-art tools. Our SCMS tool will be used to evaluate the complexity metrics, primarily following a static analysis, of any program written in Java language. The strength of this tool is that it is implemented to compute metrics at three different levels (statement, method, class). The tool outputs a CSV file of either

*Issar Arab & Email: issar.arab@gmail.com

the 20 metrics of all classes or the 9 metrics of all methods of a given Java project, based on statement level metrics. The set of metrics is discussed in later sections.

2. Related Work

Numerous tools have been implemented to compute metrics from a program's source code. These tools are very important factors in the reliability of the metrics, as their rules are well defined formulas producing after each run the same value on a similar input. There are numerous commercial, non-commercial and open source metric tools that are available for measuring code complexity.

One commonly used metric tool is an Eclipse plug-in for Java programming language maintained by IBM, which calculates numerous metrics for the source code during build cycles and warns the tester, via the problem view of 'range violations' for each metric [4]. These violations are derived from rules built in the tool which can be modified to fit local conditions. This allows the user to stay continuously aware of the program's source code's health. One of the advantages of using this Eclipse plug-in tool is the possibility of exporting the metrics in HTML, CSV or XML format.

VizzAnalyzer Framework [5] is another software tool that measures the quality of a Java implemented software system. In the VizzAnalyzer implementation, Welf et al. [5] rely on the RECODER meta-programming library [6], providing a compiler front-end for Java programs and an API for accessing AST and semantic analysis results. The software computes a set of metrics that are well established in the literature, such as metrics from the C&K metric suite as well as newly developed metrics including Data Abstraction Coupling (DAC), Package Data Abstract Coupling (PDAC), Edge Size, Tight Class Cohesion (TCC), Tight Package Cohesion (TPC), and Lack of Documentation (LOD) [5].

In what concerns other programing languages, Pymetrics is a famous open source complexity metrics measurement tool implemented by Charney [7]. It is a tool that compiles metrics for Python source code only. Metrics include McCabe's Cyclomatic Complexity [1, 8], LOC, etc. One of the main features of this tool is the output format. It allows users to customize their own reports as well as the option to output either to stdout, SQL files, or CSV files.

According to the article in [9], test tools can be divided into seven groups: design, implementation, evaluation, load, performance, GUI, dynamic and static analysis. In our study, we focus only on the last category, that deals with static and dynamic complexity metric tools. In this category, the most famous and widely used assessment metrics by far are the LOC (lines of code) and Cyclomatic Complexity [1, 8, 3]. However, the use of Object-Oriented (OO) languages has pushed researchers to come up with novel taxonomies of complexity metrics to evaluate OO solutions. The metrics introduced in [10] are very popular in dealing with Object-Oriented design and are used in the coding phase. The approach presented is based on mathematical formula that describe the relationships between the different variables. Moreover, the metrics proposed are designed to determine quantitative measures by studying the operators and operands used in the code. These metrics are referred to as 'Software Science' and are used in the development stage in order to assess the code. Another set of metrics is presented by Chidamber et al. in [11]. These metrics deal with many principles related to Object-Oriented development for the sake of enhancing software maintenance.

Numerous metrics have been introduced, discussed, scrutinized, and published. Those measurements were implemented in multiple commercial and open source tools, which raises one important question: do these metric tools produce the same values for the same metrics on the same input? This has been investigated by Lincke et al. in [12]. The authors conducted an experiment with ten commercial and free metrics tools. Using these tools, a set of nine complexity metrics were computed on the same software systems. Their investigation showed that, given the same input, metrics tools deliver different results. One reason for the disparity is the looseness of the definition of most metrics. Different tools may interpret even a simple metric such as LOC differently (e.g. are blank lines counted or not).

3. Novel Taxonomy of Complexity Metrics

In general, a metric is defined as a measurement and any measurement can be a useful metric. Software engineers use measurement throughout the entire development cycle by measuring the characteristics of software to get some notion of whether the software fulfills the requirements consistently and completely. Additionally, metrics measure the design quality and whether the software is ready for testing. Project Managers measure attributes of the product to be able to tell when the software will be ready for delivery and whether the budget will be exceeded. Customers measure aspects of the final product to determine if it meets their requirements and if its quality is sufficient. FInally, maintainers must be able to assess and evaluate the product to see what should be upgraded and improved [3].

Software metrics can be clustered into four main categories [1]:

- Product
- Process
- People
- Value of the Customer

In this paper, our focus was on the product metrics as being the main selectors for test cases. As opposed to the previous work done in the domain, we focus on the development of a comprehensive taxonomy based on two main criteria:

1. Which product level is the metric applied on?
2. And which feature does the metric measures at a given product complexity level?

The suggested taxonomy of metrics can be projected into a 2-dimensional space where each axis represents one criterion as shown in Figure 1 [3]. Each datapoint in the graph represents the metric computed by SCMS tool given its coordinates, where x-axis (kind) represents the type of metrics measured and the y-axis (scope) represents the product level at which the metric is applied. The grouping of data points in the graph visualizes the number of metrics under a similar type compiled at a given product level. For instance, the top right cluster of points shows that 4 metrics under data usage type are computed at the class level.

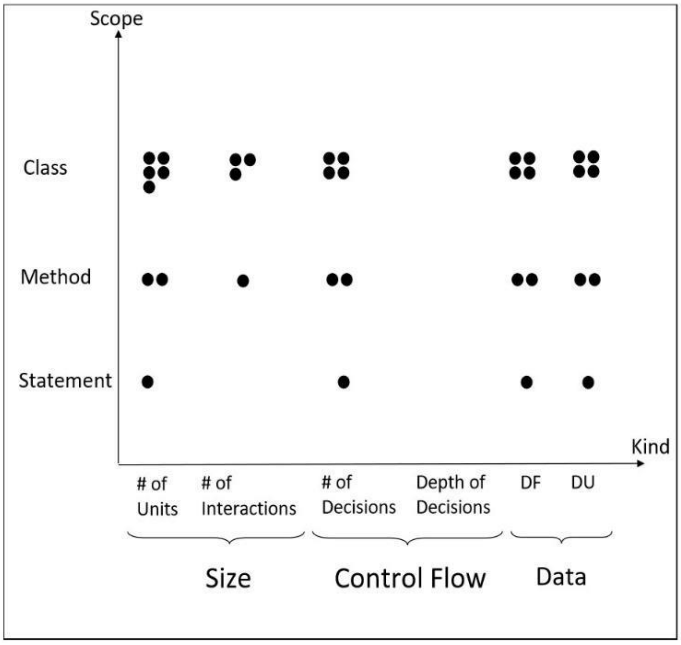


Figure 1: Our suggested taxonomy dimensions compiled by SCMS tool where y-axis (scope) represents the product level at which the metric is applied and the x-axis (kind) represents the type of metrics measured. The small clusters of datapoints in the graph shows the number of metrics under a similar type compiled at a given product level

4. Metrics at Statement Level

4.1. Data Complexity

Data complexity metrics are studied in our work by two different means, data flow and data usage. Data flow measures how the data behaves taking into consideration its interactions within the system [3, 13, 14]. It is defined as the number of 'formal parameters of activities and the mappings between activities' data. It tracks the flow of the data through the system in order to detect errors related to the usage and interactions of variables between each other. Data flow has been the choice of many testers due to its close relationship with the Object-Oriented principle of cohesion. On the other hand, data usage for a statement is defined to be the number of variable values used in that statement plus one if the computed equation is being assigned to a variable in that statement. It is based on the number of data defined in the unit being considered or the number of data related to that unit. From a software testing point of view, the data usage metric is simply the sum of uses (p-uses and c-uses) plus definition variables in a given statement.

4.2. Control Flow Complexity

McCabe's Cyclomatic complexity metric [a, h, c] is a widely and well-established measurement in the literature used in this category. It has been implemented by many software complexity tools for different programming languages and especially for Object-Oriented ones. Basically, the idea behind Thomas McCabe's metric is that software complexity increases with the number of control paths generated by its code [8]. A control path is determined by the existence of a control block, also named scope, that includes predicates as a condition to fork a flow in a path. From these observations, we chose the scope level metric of a given statement as the measurement of choice for this category. This method requires counting how many control constructs (do-while, if-else, while, for, etc.) are present in the source code [15].

4.3. Size Complexity

Size Complexity is one of the oldest measures in software engineering to quantify the length, functionality, and complexity of software [16]. One of the most widely used and easiest to compute measures in this category is the number of Lines of Code (LOC). However, many other size metrics are based on Halstead Software Science Definition [10]. It is a traditional measurement of complexity metrics tackling the issue from one single view, which is the number of operators.

Halstead measure counts all the types of operators including traditional operators (ex. +, /, and ||) and punctuations (ex. ; and ()), where one opening and closing pair of parenthesis is counted as one. From our point of view, the most important metric to identify the complexity of a statement is just the number of traditional operators. While, you can blow up Halstead metric by adding as many parentheses as one wants without altering the complexity of the system. Hence, for simplicity, we suggest using a subset of Halstead measurement dealing only with traditional operators, by counting the number of operators used in each statement.

4.4. Statement Level Metrics Summary

Table 1 represents the four metrics compiled by SCMS tool at the statement level of a Java code.

Table 1: Statement level spectrum of metrics computed by SCMS tool

Metric Short Name	Short Description
NumOp	Number of Operators
NumLev	Number of Levels
DF	Data Flow
DU	Data Usage

5. Metrics at Method Level

A method consists of a set of statements performing a specific functionality. Therefore, metrics at this granularity can be derived from previously computed statement metrics. For method level measurements, we suggest two strategies: the first is to calculate the sum of the same precomputed measurement of all statements in the method, and the second is to compute the maximum value of the same metric among all statements constructing the method [3, 17, 18]. This adds up to eight metrics at the method level derived from the statement level metrics. Additionally, we added a 9th metric, InMetCall, which counts the number of methods within the same class calling the method being studied.

Table 2 represents and defines the nine metrics compiled by SCMS tool at the method level of a Java code. The output is in CSV format including the class path to the method as well as its type.

Table 2: Method level spectrum of metrics computed by SCMS tool

Metric Short Name	Short Description
MaxOp	Maximum number of operators of each statement in the method. Metric based on the statements results
TotOp	Total number of operators of each statement in the method. Metric based on the statements results
MaxLev	Maximum number levels at each statement in the method. Metric based on the statements results
TotLev	Total number of levels at each statement in the method. Metric based on the statements results
MaxDF	Maximum number of data flow of each statement in the method. Metric based on the statements results
TotDF	Total number of data flow of each statement in the method. Metric based on the statements results
MaxDU	Maximum number of data usage of each statement in the method. Metric based on the statements results
TotDU	Total number of data usage of each statement in the method. Metric based on the statements results
InMetCall	Number of within class method calls of the method in question

6. Metrics at Class Level

The main entity in Object-Oriented design is the class. A class is a combination of attributes and methods. From this definition, we can derive the class level metrics given precomputed measurements at method level. Following the same approach applied to compute method level metrics, we again compute the total and the maximum of each measurement of the methods in a class [3, 17, 18].

We then add two additional metrics: the in-out degree of that class, which is the number of methods outside of that class that are called by at least one method in that class, and the number of public members within the class, i.e. attributes and methods in a class [c]. The public members within a class are defined as the public fields and the public methods declared in that class. Table 3 and Figure 2 depict the 20 metrics compiled by SCMS tool at the class level of a Java program as outputted in the CSV file by SCMS tool. Figure 2 shows how the 18 class level metrics, blue, are derived from the 9 method level ones, green, adding on top the In-Out Degree and the Number of Public Members.

7. Spectra Complexity Metric System (SCMS)

Spectra Complexity Metric System (SCMS), is a software that allows any Java software engineer to assess the complexity of a Java program using the above described metrics.

The tool is implemented using the Java 1.8 and Maven 3 framework. It makes use of the ASTParser pre-built Java library that allows you to convert a given Java class into an Abstract Syntax Tree for evaluation.

Table 3: Class level spectrum of metrics computed by SCMS tool

Metric Short Name	Short Description
Tot2Op	Counts the total number of operators. (method output based)
TotMaxOp	Counts the total of the max operators. (method output based)
Max2Op	Counts the max of max operators. (method output based)
MaxTotOp	Counts the max of the total number of operators. (method output based)
Tot2Lev	Counts the total number of levels in the whole class code. (method output based)
TotMaxLev	Counts the sum of the maximum level in each method.
MaxTotLev	Counts the max of the total number of levels in each method.
Max2Lev	Counts the max level in the whole class, i.e. the deepest branch. (method output based)
Tot2DU	Counts the total number of data usage in the class. (method output based)
TotMaxDU	Counts the total number of the max data usage in the class. (method output based)
MaxTotDU	Counts the max of the total number of data usage in each method. (method output based)
Max2DU	Counts the max of max data usage. (method output based)
Tot2DF	Counts the total number of data flows in a class. (method output based)
TotMaxDF	Counts the total of the max data flows in each method of the class. (method output based)
Max2DF	Counts the max of max data flows in each method of the class. (method output based)
MaxTotDF	Counts the max of the total data flows in each method of the class. (method output based)
TotInMetCall	Counts the total number of within class method calls. (method output based)
MaxInMetCall	Counts the max number of within class method calls. (method output based)
inOutDeg	Counts the number of in class call of external methods. Similar to out degrees of a dynamic call graph.
pubMembers	Counts the number of members in a class.

Figure 3 shows the UML class diagram of our tool, where CMSRunner class represents the main entry point of the software. The main class uses FileUtils functionalities to extract all the java files in the project to be assessed; then, it translates each file into an abstract syntax tree via the class Parser. This parse tree is then traversed, using AstClassExplorer and AstStatementExplorer, to retrieve the relevant information to our suggested taxonomy in

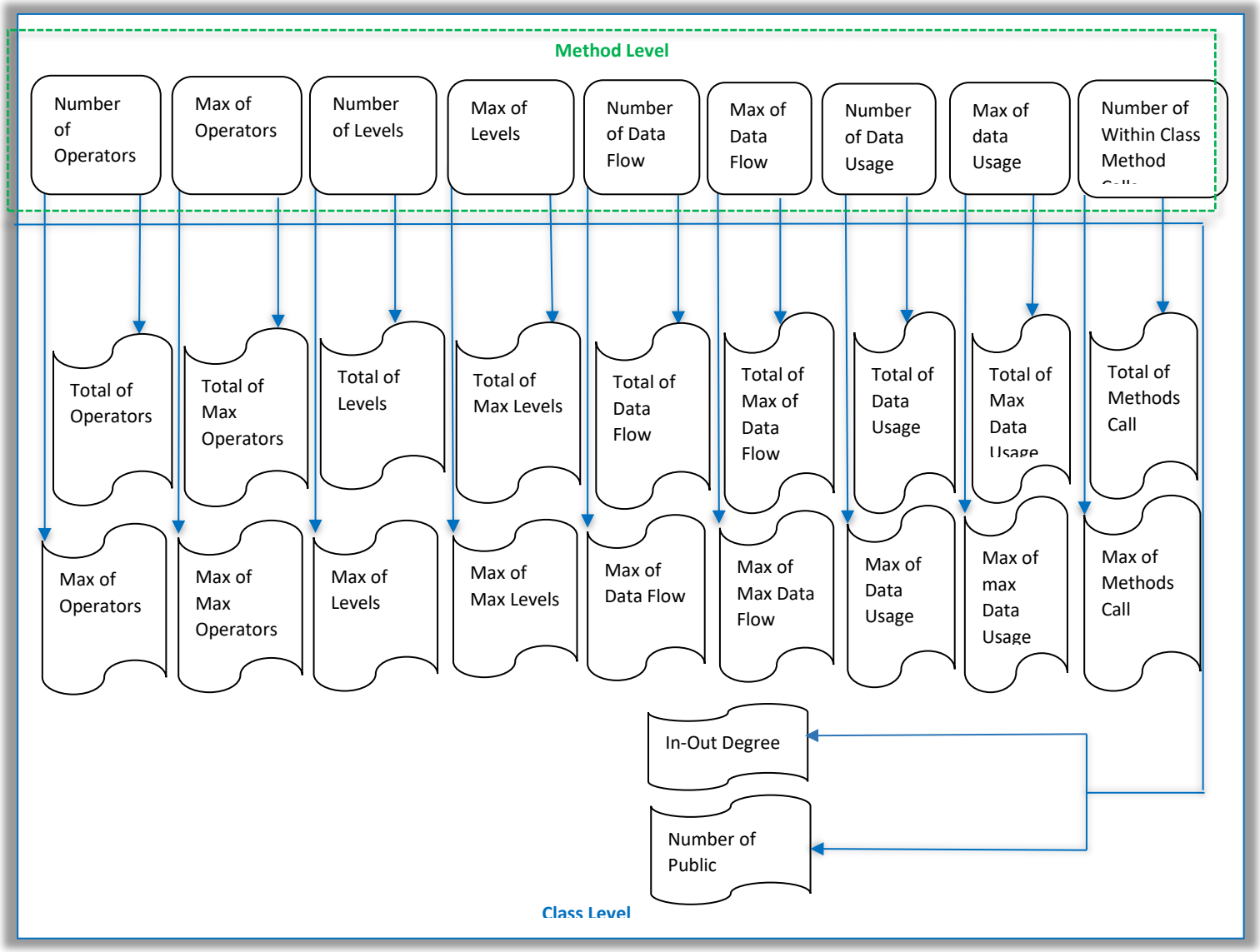


Figure 2: Taxonomy of class level complexity metrics derived from the method level ones as they are compiled by SCMS tool

order to populate the metrics fields. The metrics are organized by Scope (product level) as fields in the classes Statement, Method, and Class shown in the UML diagram.

8. SCMS Tool Evaluation and Benchmarking

In this evaluation, we focus on the product metrics category to compare existing software tools with our suggested taxonomy. As a first step, we want to showcase the importance of automated tools to calculate complexity metrics. Then provide a detailed description of the selected tools in order to make the comparison with our software. And finally, we compared our SCMS tool with the different tools based on three main levels, namely statement, method, and class level.

8.1. Manual vs. Automated Metric Evaluation

Software quality assessment has been researched by engineers and pioneers in the field to express quality attributes of a software. Therefore, solid numbers and measures should be generated for a program to better assess its complexity. Many tools have been developed to collect metrics from a system that has been implemented. Moreover, these tools have differences in myriad aspects.

The two questions that might be asked here are: why do we need tools? And why can't we measure the complexity of a program manually or only by reading through the program's code? The answer to these questions is that the complexity of a software has increased tremendously [16]. A system now may contain thousands of lines of code which makes measuring its complexity manually not viable.

Therefore, there is a need for tools that analyze files and components of a system rapidly and generate reports. The results will then be used to optimize the process of test case generation. These measurements can give a software developer a deep view of the different non-functional requirements such as maintainability and performance. Furthermore, a tool follows repeatable and consistent calculation procedures in assessing a

program. Therefore, compiled information can be easily interpreted and analyzed via graphical and data visualization techniques. On the other hand, calculating metrics manually adds an overhead to the software testers, which may lengthen the testing phase,usually taking much more time than the development period. The manual process is very tedious, especially when it comes to complex software that contains a very large number of units [16]. Therefore, complexity metrics tools are essential in assessing the quality of a program.

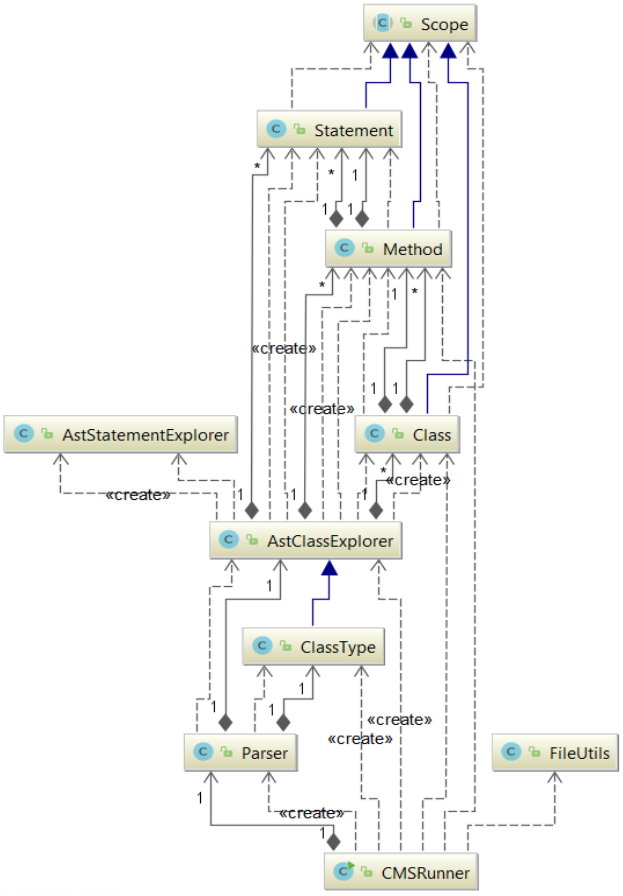


Figure 3: Simplified UML class diagram of SCMS tool

8.2. Selection Process of Tools

An investigation through the available tools online showed that most of the tools provide process and product quality attributes. In order to make a reliable benchmarking of our software, we only picked tools computing product quality metrics based on three criteria. The first criterion was to select tools that have a large set of product metrics. Second, the focus was only on tools calculating complexity metrics for object-oriented programming languages. Third, the selection was based on the ease and regularity of how tools are deployed and used. Usability of the tools was a key selection criterion in our study.

8.3. Selected Tools Summary

8.3.1 JHawk Tool

The first software tool we examined is called JHawk. JHawk was developed to compute complexity metrics for Java programs. The tool is described mainly as a static code analysis tool since it takes as input the program code to compute metrics. The computed metrics are based on several aspects ranging from relationships between components, volume, and complexity. The JHawk tool delivers both the process and the product quality metrics. We retrieved only the list of product metrics. JHawk calculates complexity metrics on all levels. The system level is the container or the upper layer as it wraps all the remaining levels which are method, class, and package level. Table 4 shows a sample set of metrics calculated by the JHawk tool at the method and class levels [19].

Table 4: Class and method level metrics measured by JHawk

Level	Metric Name	Metric Description
Method	Number of Arguments	The number of arguments per method
	Variable Declaration	The number of variables declared in the method
	Number of Statements	The number of statements in the method
	Number of Loops	The number of loops (for, while…) in the method
	Number of Operators	The total number of operators in the method
	Number of Operands	The total number of operands in the method
Class	Number of Statements	The number of statements in the class
	External Method Calls	The external methods called by the class and methods in the class
	Lines of Code	The number of lines of code in the class and its methods
	Modifiers	The modifiers (public, protected, etc) applied to the declaration of the class
	Local Methods Calls	The local methods called by the class and by methods in the class that are defined in the hierarchy of the class
	Instance variable	The instance variables defined by the class

8.3.2 Analyst4J

Analyst4J was developed both as a plugin for Eclipse Integrated Development Environment, and as an application that can be installed and executed independently [12]. Analyst4J was designed to provide several functionalities of a Java program including:

- Computing metrics
- Analysing source code
- Generating reports containing information about a Java program

Analyst4j V1.5.0 version supports regression testing by analyzing code before and after being changed. It provides an easy to read report about the design and maintainability of the system to be used. Moreover, code quality can be easily analyzed by the tool. Analyst4J can be integrated with bug finding tools in order to maintain Java programs. It also provides several complexity metrics for java programs including complex classes and methods.

8.3.3 CCCC Tool

CCCC is a software tool designed mainly to measure Cyclomatic Complexity of code in a program. It is generally used for two main programming languages which are C and C++. The generated output from this tool is a web page containing detailed information of all the program source code. This tool is freely available for download and use [20].

CCCC generates the following information about a given program:

- The number of lines for each module in a program
- McCabe Cyclomatic Complexity metric
- Coupling between different components
- Object-oriented metrics as depth of inheritance tree
- Chidamber and Kemerer complexity suite

8.3.4 JMetric Tool

This tool was created by Cain and Vasa [21] for a project at COTAR in Australia. It calculates object-oriented metrics for the Java code. JMetric was installed to retrieve the product metrics it calculates. This tool computes complexity metrics at five different levels: project, package, class, method, and variable level. Table 5 shows a sample set of the computed metrics by JMetric at the method and class levels relevant to our product measurements comparison study.

Table 5: Class and method level metrics measured by JMetric

Level	Metric Name	Metric Description
Method	Number of Statements	The number of statements in the method
	Collaboration	Collaboration with other methods or units
	Cyclomatic Complexity	Cyclomatic complexity of the method
	Lines of Code	The total number of lines of code in the method
Class	Number of Statements	The number of statements in the class
	Number of Methods	The number of methods in the class
	Lines of Code	The number of lines of code in the class and its methods
	Public Methods	Count the number of public methods in the class
	Local Methods Calls	The local methods called by the class and by methods in the class that are defined in the hierarchy of the class
	Instance variable	The instance variables defined by the class

8.3.5 NDepend Tool

NDepend was designed to analyze code developed in .NET technology. This tool has a large support for code metrics since it analyzes the program code at six different levels:

- Metrics on application
- Metrics on assemblies
- Metrics on namespaces
- Metrics on types
- Metrics on methods
- And metrics on fields

Table 6 shows a partial list of metrics calculated by the NDepend tool [22].

Table 6: Sample of complexity metrics computed by NDepend

Level	Metric Name	Metric Description
Application	NbLinesOfCode	Computing the number of logical lines of code of the application
	NbLinesOf Comment	Number of lines of comment that can be found in the application
	NbMethods	Number of methods for the whole application/program
	NbFields	The number of fields. The field can be either regular, enumeration value, or other types.
Method	NbLinesOfCode	Number of lines of code inside a method
	NbOverloads	The number of overloads of a method
	NbLinesOf Comment	Number of comments within a method
	NbParameters,	Number of methods inside a method
	NbVariables	Number of variables in a method
Field	Size of Instance	The size in bytes of instances
	Afferent coupling at field level (FieldCa)	The number of methods that directly use a field

8.3.6 Chidamber and Kemerer Java Metrics (CKJM) Tool

The Chidamber and Kemerer metrics were among the first introduced metrics in the history of software metrics.

CK metrics objectives are [23]:

- To measure unique aspects of OO approach
- To measure complexity of the design
- To improve the development of the software

Diomidis Spinellis developed a tool named *CKJM* [24] to compute the Chidamber and Kemeres suite of metrics. This tool measures these metrics by processing the bytecode of compiled Java files. Java files are compiled prior to be given as input to the *CKJM*. For each class of the program or project, *CKJM* provides as output the six well known metrics of Chidamber and Kemerer. The six metrics that are calculated by the *CKJM* are:

- WMC: Weighted methods per class
- DIT: Depth of Inheritance Tree
- NOC: Number of Children
- CBO: Coupling between object classes
- RFC: Response for a Class
- LCOM: Lack of cohesion in methods

8.4. Comparison to SCMS tool

In our benchmarking, we considered only similar metrics to our presented novel taxonomy computed by CSMS and discarded the rest. The four stated metrics at the level of a statement will be used as the basis to derive other high-level complexity metrics, namely for the method and class level.

Table 7: Statement level comparison

Tools	Metrics			
Name	Level Number	Data Flow	Data Usage	Number of Operators
SCMS	**x**	**x**	**x**	**x**
Analyst4j		x		
JMetric		x		
CCCC		x		
CKJM		x		
JHawk		x		
NDepend		x		

Table 8: Method level comparison

Tools	Metrics								
Name	Total # of levels	Max # of levels	Total # of operators	Max # of operators	Total # of DF	Max # of DF	Tot # of DU	Max # of DU	# In Method Calls
SCMS	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**
NDepend	x	x	x		x	x			
Analyst4j	x		x	x	x	x			
JMetric			x	x	x	x			
CCCC			x		x	x			
CKJM			x		x	x			
JHawk			x		x	x			

As illustrated in Table 7, the only metric that is measured at the statement level by all the selected tools is the data flow (DF) metric. This shows the uniqueness and the added value of our tool. Likewise, at the method level in Table 8, we see that the closest tools to our software are NDepend and Analyst4j with a coverage of 55% of our suggested taxonomy. Also, three metrics in our suggested set is covered by all of the selected tools. Those are the total number of operators, the total number of DF and the maximum DF. Furthermore, since the class level metrics are derived from the method level ones by computing the sum and the maximum, we can conclude that only six metrics from a total of 18 metrics are covered at the class level by all those selected tools. Therefore, we show the novelty and the added value of our suggested taxonomy and Spectra Complexity Metric System.

9. Conclusion and Future Work

This paper presents a tool, SCMS, which evaluates and assesses a novel comprehensive taxonomy of complexity metrics, following a bottom-up approach, of any given software written in Java language. These measurements can then be used to optimize the process of test case generation by targeting the complex units in the system. The taxonomy tackles the issue from two dimensions. The first being the scope/product dimension and it covers the metrics at three granularity levels: statement, method, and class level. The second being the type of evaluation, which covers three types of metrics measurements: the size focusing on the number of units/operators, the control flow focusing on the number of decisions, and the data type focusing on both the data flow and the data usage. Those metrics have as a goal to increase the efficiency of the testing process by significantly reducing the number of test cases without having a significant drop in test effectiveness. The paper also covers the design and technology used to implement SCMS tool by providing a UML class diagram of the software which can offer insights for a starting point of any future development of new complexity metric tools. Finally, we have compared our tool with the major existing Java related tools that cover a large set of product level metrics in the market. We showed the novelty of our taxonomy of complexity metrics and the capability of our tool to compute these measurements based on the three different product categories. The tool is available as an open source at https://github.com/issararab/SCMS .

In the future, we will investigate the usability of this tool as well as other tools with their potential in software diagnosability and fault detection. Our aim is to gather a large set of features/metrics and combining them with machine learning algorithms to artificially predict the faulty units in a software. Our unit level of interest as of now is the class level.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to thank reviewers for their exceptionally constructive reviews to improve this work and inspire future work. The authors would like also to thank their sponsor, Al Akhawayn University in Ifrane, for funding this project.

References

[1] T. J. McCabe, C. W. Butler, “Design complexity measurement and testing”, in 1989 ACM 32, 12, 1415-1425. DOI: http://dx.doi.org/10.1145/76380.76382

[2] I. Arab, S. Bourhnane, Reducing the cost of mutation operators through a novel taxonomy: application on scripting languages, in 2018 Proceedings of the International Conference on Geoinformatics and Data Analysis (ICGDA '18). ACM, New York, NY, USA, 47-56. DOI: https://doi.org/10.1145/3220228.3220264

[3] F. Bouchaib and K. Magel. (2015). “Taxonomy Dimensions of Complexity Metrics” Int'l Conf. Software Eng. Research and Practice , pp. 96-102: Las Vegas, Nevada, USA. Available: https://www.researchgate.net/publication/281858313_Taxonomy_Dimensions_of_Complexity_Metrics

[4] https://marketplace.eclipse.org/content/eclipse-metrics (accessed on May 2019)

[5] L. Welf, E. Morgan, L. Jonas, P. Thomas, P. Niklas, VizzAnalyzer- A Software Comprehension Framework, 2003 Available: https://www.researchgate.net/publication/241753972_VizzAnalyzer-_A_Software_Comprehension_Framework

[6] A. Ludwig, RECODER Homepage, http://recoder.sf.net, 2001.

[7] R. Charney, Programming Tools: Code Complexity Metrics, Jan 2005, Available: https://www.linuxjournal.com/article/8035

[8] T. J. McCabe, “A Complexity Measure”, in July 1976 IEEE Trans. Softw. Eng. 2, 4, 308-320. DOI: https://doi.org/10.1109/TSE.1976.233837

[9] P. Pohjolainen, Software Testing Tools, Department of Computer Science and Applied Mathematics, 2002. Available: http://cs.uef.fi/uku/tutkimus/Teho/SoftwareTestingTools.pdf

[10] M. H. Halstead. "Elements of Software Science (Operating and Programming Systems Series)". Elsevier Science Inc., New York, NY, USA, 1977

[11] S. R. Chidamber, C. F. Kemerer, "A Metrics Suite for Object Oriented Design". in June 1994 IEEE Trans. Softw. Eng. 20, 6, 476-493. DOI=http://dx.doi.org/10.1109/32.295895

[12] R. Lincke, J. Lundberg, W. Löwe,"Comparing software metrics tools". in 2008 In Proceedings of the 2008 international symposium on Software testing and analysis (ISSTA '08). ACM, New York, NY, USA, 131-142. DOI: https://doi.org/10.1145/1390630.1390648

[13] J. Cardoso, “Control-Flow Complexity Measurement of Processes and Weyuker’s Properties,” Word Academy of Science, Engineering and Technology, August 2005.

[14] J. Cardoso, “Control-flow Complexity Measurement of Processes and Weyuker's Properties.” in 2007 International Journal of Mathematical and Computational Sciences Vol:1, No:8, 2007. doi.org/10.5281/zenodo.1076476

[15] J. Cardoso, How to Measure the Control-flow Complexity of Web processes and Workflows, Workflow Handbook, Future Strategies Inc., pp.199-212, 2005

[16] A. Arenas, J. Duch, A. Fernandez, S. Gomez, “Size Reduction of Complex Networks Preserving Modularity”, in 2007 New Journal of Physics 9, P. 176, iopscience.iop.org/article/10.1088/1367-2630/9/6/176

[17] Y. Shin, A. Meneely, L. Williams, J. A. Osborne, "Evaluating complexity code churn and developer activity metrics as indicators of software vulnerabilities", in 2011 IEEE Transactions on Software Engineering, vol. 37, no. 6, pp. 772-787

[18] B. Falah, K. Magel. (2013) “Test Case Selection Based on a Spectrum of Complexity Metrics”. Proceedings of 2012 on International Conference on Information Technology and Software Engineering (ITSE), Lecture Notes in Electrical Engineering , Volume 212, pp. 223-235, DOI: 10.1007/978-3-642-34531-9_24

[19] Virtual Machinery, "Download JHawk Trial Version", Virtual Machinery, 2017. [on ligne]. Available: http://www.virtualmachinery.com/jhdownload.htm

[20] B. Wicht, (August 2011) “How to Compute Metrics of C++ project using CCCC”. [on ligne]. Available: https://baptiste-wicht.com/posts/2011/08/compute-metrics-of-c-project-using-cccc.html

[21] A. Cain, R. Vasa, and D. Franek, Source Forge.(accessed. 20.05.2019) Available: https://sourceforge.net/p/jmetric/wiki/Home/

[22] P. Smacchia, “Partitioning Your Code Base Through .NET Assemblies and Visual Studio Projects,” Redgate Hub, accessed. 10.06.2019 Available: https://www.red-gate.com/simple-talk/dotnet/net-framework/partitioning-your-code-base-through-net-assemblies-and-visual-studio-projects/.

[23] S. R. Chidamber and C. F. Kemerer , “Towards a Metrics Suite for Object Oriented Design”, in October 1991 C onference proceedings on Object-oriented programming systems, languages, and applications (OOPSLA '91), Andreas Paepcke (Ed.). ACM, New York, NY, USA, 197-211. DOI=http://dx.doi.org/10.1145/117954.117970

[24] Diomidis Spinellis. Tool writing: A forgotten art? IEEE Software, 22(4):9–11, July/August 2005. (doi:10.1109/MS.2005.111).

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 222-225 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Stability of Ionized Air Using Density Functional Theory (DFT) for High Voltage Engineering

Fri Murdiya[1], Neni Frimayanti[2,*]

[1]*Department of Electrical Engineering Faculty of Engineering Universitas Riau Kampus Bina Widya, Jl. HR, Subantas KM 12.5 Simpang Baru Tampan Riau Indonesia*

[2] *Sekolah Tinggi Ilmu Farmasi (STIFAR) Riau, Jalan Kamboja, Pekanbaru, 28293, Indonesia*

ARTICLE INFO

Article history:
Received: 24 October, 2019
Accepted: 25 November, 2019
Online: 22 January, 2020

Keywords:
DFT
Ionization
Exitation energy
High voltage
Gas insulation

ABSTRACT

In high voltage engineering, air is the one insulation gas that can ionized cause by electrical discharge. Air consisted of many gases such as O2, N2, CO2, H2, etc. In order to study the characteristic of the insulation gas, we can apply density functional theory (DFT) to determine gas ionization level such as charge distribution and excitation energy. These data are shown the stability of gas molecule when ionization process occure cause by electrical discharge, ultraviolet light, etc. For application on ozone production by electrical discharge, the oxygen atom can transform to ozone need energy above the stabiity excitation energy of oxygen.

1. Introduction

In recent year, the high voltage engineering is focused on electrical discharge on the insulation dielectric in the form solid, liquid and gas. Base on our previous research, computational approach can be used to study the air failure in the high voltage which can be utilized to generate ozone from the ionization process of the oxygen atoms [1]. The ionization process has been checked using homo-lumo application. Some gases such as hydrogen, oxygen, nitrogen, carbon dioxide widely used as high voltage insulation. These gases are also used as an ionized gas for application in the high voltage like to produce plasma. This plasma was then used for physics, chemistry and medicine fileds. Recently, laboratory testing was required to determine the electrical strength (E) of that gas [2] or by kinetic modeling of the breakdown process on the basis of a microscopic electron-molecular cross section [3]. These methods are needed much more time and resource intensive and only can be applied to narrow gas.

There are some obvious advantages of the calculation method compared than traditional methods such as saving cost in term of time and effort. Moreover the calculation approach allows for identifying the atomistic processes of insulation gas, and also for exploring the intrinsic changes inside the sensing materials. Along with the development of computational materials science, the theoretical calculation methods can not only be used to provide the reasonable and profound explanations for the experimental results, but they are helpful for the design of new materials structures and thus new functionalities Previous reports reported that E is highly correlated with certain predictors which are simple functions of certain molecular properties [4]. These molecular properties are calculated using Density Functional Theory (DFT). Thus, E for molecules can be predicted by their own computational methods before testing their electrical and chemical behavior in experiments.

In this research, density functional theory (DFT) has been developed as reliable and effective approach for the computer simulation of molecular structure and also to measure the energy for chemical reaction. DFT is also can be used to search the efficiency and acuracy of the molecular properties, such as electronegativity, chemical potential and affinities [5]. The molecular properties of can be accurately described and calculated by manipulating the electron density and its fundamental quantity.

2. Research Methodology

2.1. Preparation of Molecular Gases

Preparation of molecular gases i.e. O2, N2¸ H2, CO and CO2 are beginned with skecthed of each molecular gases using ChemBioDraw ultra 13.0 for then it should save it in mdl. mol format. for then imported it into the Gaussian software.Finally, complete content and organizational editing before formatting.

*Corresponding Author Neni Frimayanti, Address : STIFAR Riau, Jalan Kamboja, Pekanbaru, Indonesia,, email: neninfrimayamnti@gmail.com

www.astesj.com
https://dx.doi.org/10.25046/aj050128

Please take note of the following items when proofreading spelling and grammar.

2.2. Density Functional Theory (DFT)

The initial structure of each gas O2, N2, H2, CO and CO2 are needed to obtain, density functional theory (DFT) is urgently needed to optimize the stability of these gases. DFT was runned using Gaussian 09 software with 6-311G/B3YLP was selected as basis set. Basis set is shape of atomic orbital that explained using wave function. TD-SCF was used as method to perform the Gaussian [6]. Stability of the calculation was explored the stability of the wavefunction computed for a molecular system. The stability of SCF solution for unknown systems should always be tested. The consideration for stability of unknown systems should be apply and may be tested in calculation using DFT methods as well.

3. Results and Discussion

Single particle are described about molecules or atom or also described how the chemical bonding between each atom to make a molecule. In this study, density functional theory (DFT) was applied to check the stability of gases and also to study shock induced process because they involve spatial and temporal scales that go beyond those attainable at the atomistic level. The stability of each gas such as N2¸O2, CO, CO2¸ and H2. base on the DFT calcution we will know charge distribution and excitation energy [7]. DFT calculation provides the energy of ionization potential for these gas elements.

Based on DFT simulation for these gases N2, O2, H2, CO and CO2, these gases are looks stable with the excititation energy of 0.56334 eV, 0.56334 eV, 0.49436 eV, 0.52185 eV, 0.53651 eV, respectively. Plate testing model (anode and cathode electrodes) was used for high voltage for then it can be simulated the excitation energy using the equation 1 (eq. 1) with d = 0.01 m. The results of the calculation of excitation energis is presented in Table 1.

$$W = F.d \quad (1)$$

Where, W is excitation energy (eV) and F is Coulomb force on electrical charges (N)

$$F = E \,.\, e \quad (2)$$

E is electric Field/ electrical strength (V/m) and e is electron

Initially, the DFT for each gas was evaluated using Gaussian 09 software package. Instead of DFT, it also can predict the spectrum UV-Vis for each gas. Based on spectrum UV-Vis, the information about energy excitation will be informed [8].

3.1. Nitrogen

The gas N_2 is seems to be stable at the wave length 89.17 nm, based on DFT calculation for the best pose selected with the following results; energy of excitation of 89.17 eV, energy for the ionization process of 108.98 au from orbital eight of -0.13492 to orbital seven of -0.42842. The ionization process of molecule N_2 is depicted in Figure 1.

Molecule consisted of atoms, atoms consisted of electrons, proton and neutron. Excitation of each electron in orbital atoms can cause the ionization process. The ionization process is stable if the range energy between two orbital less than 0.9 eV [9]. For gas N_2, there are two electrons was jumped from orbital eight (4p) into orbital seven (4s) with the range energy of 0.56334 eV. It is indicated that gas N_2 is stable.

Table 1. Gases excitation energy

No	Gases	W (eV)	E (V/m)
1	Molecule N_2	0.56334	56.334
2	Molecule O_2	0.56334	56.334
3	Molecule H_2	0.49436	49.436
4	Molecule CO	0.52185	52.185

There are some reasons to make the frontier orbital become so significant, they are: method for calculating the ionization energy and Koopman's theory. Based on Homo-Lumo calculation on our previous paper [10], it is provided the energy of ionization potential for gas elements. Furthermore, DFT calculation is needed to analyze much more about the orbital for basic gas.

Each molecule has orbital or molecular orbital (MO), it can be used to represent the regions in a molecule where an electron occupying that orbital is likely to be found. Molecular orbitals also can be used to predict the location of an electron in an atom. A molecular orbital can specify the electron configuration of a molecule: the spatial distribution and energy of one (or one pair of) electron(s). The molecule will stabilize if the electrons in the last orbitals have their paired up [11].

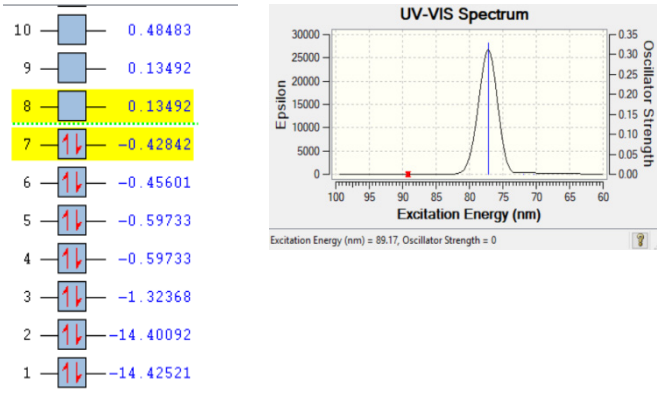


Figure 1: Ionization process of gas N_2

3.2. Oxygen

Generally, DFT can be defined as method for computational quantum mechanical modelling used in physic, chemistry or material science for investigating the properties of electron in a system, hence the name density functional theory comes from the use of functionals of the electron density [12].

Like gas N2, ionization of molecule O2 is looks stable since the excitation energy of -395143 eV with the wave length of

395143 nm, the best pose selected with the following results; energy for the ionization process of 108.98 au from orbital nine of 0.01523 to orbital eight of -0.05553. The ionization process of molecule O2 is presented in Figure 2. Stability of gas O2 is shown with there are two electrons was jumped from orbital nine (4d) into orbital eight (4p) with the range energy of 0.07053 eV.

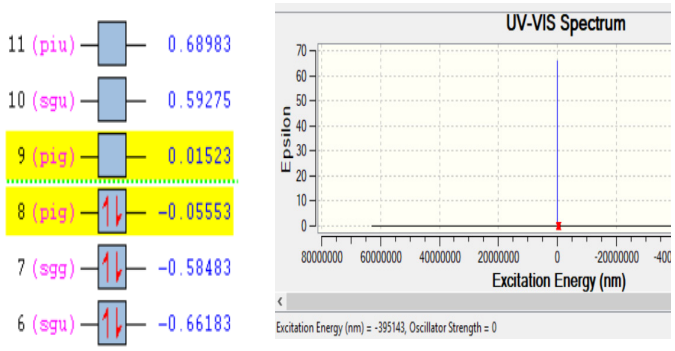


Figure 2: Ionization process of gas O_2

3.3. Hydrogen

The stability of molecule H_2 is also checked using the same way like molecule O_2 and N_2. Base on DFT calculation, it come out with the excitation energy of 88.8 eV at the wave length 88.8 nm. Two electrons were jumped from orbital two into orbital one with the energy of 0.49436 eV. Figure 3 is presented the ionization process of molecule H_2.

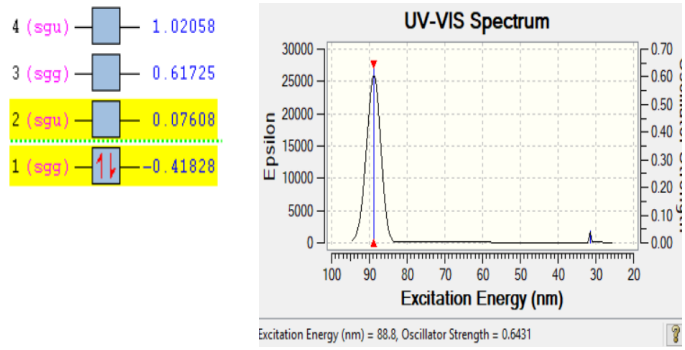


Figure 3: Ionization process of gas H_2

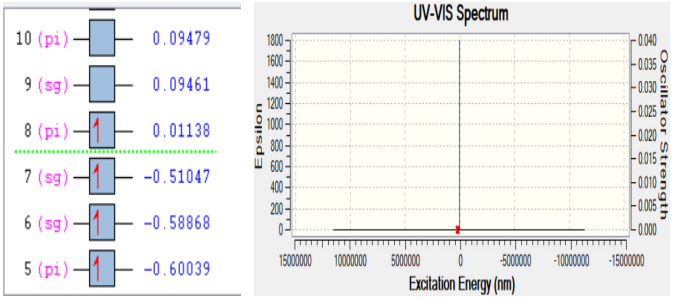


Figure 4: Ionization process of gas CO

3.4. Carbon Monoxide

Stability of molecule CO was checked using DFT. The first principles of DFT is to carried out the structural stability and also the electronic properties of a molecule. Base on DFT calculation, gas CO has the excitation energy of 167782 eV. Only one electron was excited from orbital eight into orbital seven with the energy of 0.52185 eV. The process of electron excitation for molecule CO is depicted in Figure 4.

3.5. Carbon Dioxide

DFT determined the stability of molecule CO2. Base on the DFT simulation, gas molecule CO2 is also looks stable with the excitation energy of 96.13 eV. There are two electrons were excited from orbital eleven into orbital ten with the energy of 0.53651 eV. Figure 5 is presented the ionization process of molecule CO2.

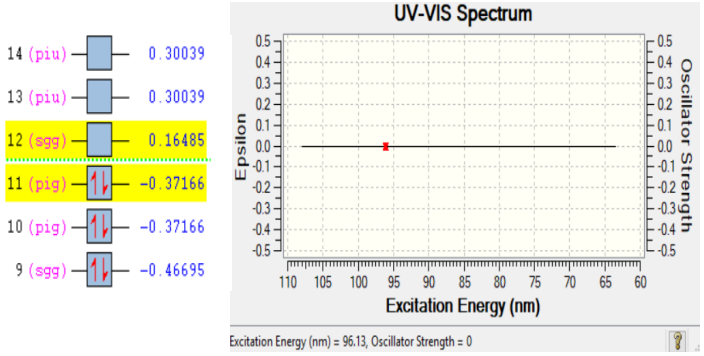


Figure 5: Ionization process of gas CO_2

4. Conclusion

There are several conclusions that can be concluded base on this research:

1. Density functional theory (DFT) has been developed as reliable and effective approach for the computer simulation of molecular structure.

2. The ionization process is stable if the range energy between two orbital less than 0.9 eV.

3. Gas molecule N_2, O_2, H_2, CO and CO_2 is stable with the excititation energy of 0.56334 eV, 0.56334 eV, 0.49436 eV, 0.52185 eV, 0.53651 eV

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

We thank to Lembaga Penelitian dan Pengabdian Kepada Masyarakat (LPPM) Universitas Riau for the financial support through Penelitian DIPA UNRI percepatan inovasi 2019, contract Number :1026/UN.19.5.1.3/PT.01.03/2019

References

[1] L. B. Affoue, B. A. Amon, D. Y.. Kicho, G. K. Mamadou, K. Soleymane, K. Mawa, R. N. Boka, S. B. El-Hadj, "Theoretical study by density functional theory method (DFT) of stability,tautomerism, reactivity and prediction of acidity of quinolein-4-one derivatives" Computational Chemistry. **6**, 57-70, 2018. https://doi.org/10.4236/cc.2018.63005

[2] G. Biasutti,'Neue Isoliergase für gekapselte Hochspannungsanlagen". Ph.D. Thesis, Diss. ETH Nr. 7683, 1985

[3] R. E. Robson, P. B. Nicoletopoulos, R. D. Li, "Kinetic theoretical and fluid modelling of plasmas and swarms: the big picture" Plasma Sources Science and Technology, **17**, 024020, 2008. https://doi:10.1088/0963-0252/17/2/024020

[4] M. Rabie, D. A Dahl, S. M. Donald, M. C. Reiher, M. Franck,. "Predictors for Gases of High Electric Strength". IEEE Transactions on Dielectrics and Electrical Insulation, **20** (3), 856–863, 2013. https://doi: 10.1109/TDEI.2013.6518955

[5] M. S. Jorge, "An introduction to density functional theory in chemistry" 2018. Theoretical and computational chemistry, **2**, 1-27, 2018. https://doi.org/10.1016/S1380-7323(05)80031-7

[6] M. J. Frisch, G. W. Trucks, H. B. Schlegel, "Gaussian 03, Revision A.1" Gaussian Inc., Pittsburgh, 2003

[7] A. Nathan, M. Guy, "Density Functional Theory — an introduction" American Journal of Physics. **68** (1), 69-79, 2000. https://doi.org/10.1119/1.19375

[8] K. Erick, L.Vladimir, "Study of gas-phase O–H bond dissociation enthalpies and ionization potentials of substituted phenols – Applicability of ab initio and DFT/B3LYP methods" Chemical Physics, **330**, 515–525, 2006. https://doi.org/10.1016/j.chemphys.2006.09.026

[9] J. Ahmad, J. Li, M. Hussain, M. T. Riaz, M. A. Akbar, and M. Z. Khan, "Study of Physio-chemical Properties of POSS/Mineral Oil based Nanofluids," in 2018 International Conference on Power Generation Systems and Renewable Energy Technologies (PGSRET), 2018, pp. 1–5. https://doi.org/10.1109/PGSRET.2018.8685972

[10] M. Kurt, T. R. Sertbakan, M. Ozduran, "An Experimental and Theoretical Study of Molecular Structure and Vibrational Spectra of 3- and 4-Pyridineboronic Acid Molecules by Density Functional Theory Calculations". Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, **70,** 664-673, 2008. https://doi.org/10.1016/j.saa.2007.08.019

[11] Q. Hai, L. Wenhao, W. Xiaoxing, C. Wei, D. Jun, H_2S and SO_2 adsorption on Pt-MoS_2 adsorbent for partial discharge elimination: A DFT study. Results in Physics, **12**, 107-112. 2019. https://doi.org/10.1016/j.rinp.2018.11.035

[12] Z. Yubo, W. James, X. F. Bing, S. Jianwei, "Subtlety of TiO_2 phase stability: Reliability of the density functional theory predictions and persistence of the self-interaction error" The Journal of Chemical Physics. **150** (01), 014105. 2019. https://doi.org/10.1063/1.5055623

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 226-232 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Framework for Automation of Cloud-Application Testing using Selenium (FACTS)

Md Nurul Islam*, Syed Mohammad Khurshaid Quadri

Department of Computer Science, Jamia Millia Islamia (A Central University), 110025, India

ARTICLE INFO

Article history:
Received: 10 December, 2019
Accepted: 08 January, 2020
Online: 22 January, 2020

Keywords:
Test
Automation
Framework
Selenium

ABSTRACT

A framework is an amalgamation of integrated tools, libraries, utilities and its coordination to interact with the other automation components. The motive of designing a test automation framework is to implement uniform standards towards automation throughout the organization to achieve the desired outcomes. Test automation framework is required to maintain the standardization of the activities performed to manipulate the operations (addition, modification and deletion) on the test scripts and functions easily, to provide the scalability and reliability. Tools, process and testing team are the three key player of test automation. The proposed framework is based on the consideration of test environment (includes language binding, IDE, automation tools and testing framework). Selenium WebDriver is used as web application automation framework to execute test across multiple browsers that supports multiple operating system with variant programming language. Selection of WebDriver automation tool for this framework is due to their internal architecture to directly communicate with browser for fast execution. The combination of test automation tools (i.e. Java, Eclipse, Selenium and TestNG) makes JEST for designing test automation framework. Both Positive and negative test must be performed to verify functionalities of applications to handle unusual exceptions.

1. Introduction

Software testing is an evaluating as well as mitigating process, as it evaluates the functionalities of application to meet out the specific requirements of software product, at the same time it also identify the defects in product to mitigate the risk of software failure by find bugs and errors in the program.

"Framework for Automation of Cloud-application Testing using Selenium (FACTS)" is almost self-explanatory, though this paper may also explore the terminology explicitly. This paper may discuss about, why do we require test automation framework, how the software testing framework may organize the test automation, how do we select best appropriate software testing framework, what are the challenges being faced by the software automation team to design and deploy the test automation framework, what are the potential benefits of using software testing framework.

Why selenium for automation? Selenium is powerful open source tool for test automation framework independent of language binding & development platform that support functional testing across the browsers.

A framework is a combination of various forms of principles, guidelines, processes, concepts and tools etc. followed by testers to take advantage of it, to automate the application. Test automation framework decides the success or failure of any software automation project. Test automation is testing process of automating test with the help of automation tools to find out defects in the application. Test automation reduces test execution time and testing cost along with the less human interaction with the system. Test script has been used to execute test multiple times, when required. Test script must consider, what to automate or what not to be, like functionalities that are going to be change in near future or not, whether script is applicable for regression testing or not, whether test cases takes more time to execute/execute with errors or operated over huge amount of input data etc.

2. Test Automation Framework – An Understanding

Test automation framework is set of guidelines/rules followed while writing test script to generate test cases as per the

* Md. Nurul Islam, Department of Computer Science, Jamia Millia Islamia, New Delhi, India 110025, 9818933615 & mnislam@jmi.ac.in

requirement of test scenarios. A framework is an amalgamation of integrated tools, libraries, utilities and its coordination to interact with the other automation components to execute the test scripts and to capture the outcomes in systematic and beneficial manner to test a specific product. If a testing activity automated without having a test automation framework, it is obvious to raise the question that "how reliable the method would be? And how quickly they can produce the same output?"

The motive of designing a test automation framework is to implement uniform standards/policies towards automation throughout the organization to achieve the desired result. A test automation framework can be considered effective if the result of test automation is easily understandable to end user, otherwise executing high volume of test cases are of little or no use for end user. The automated test generated report must include the point where the application gets failure along with test data being used for automation.

2.1. Why do we require test automation framework?

Test automation framework is required to maintain the standardization of the activities performed to manipulate the operations (addition, modification and deletion) on the test scripts and functions easily, with the aim to provide the scalability and reliability. If a tester automate testing activities without considering a framework, it is obvious to ask how reliable the method would be. Can the tester re-generate the same result? An automation framework sort out all such problems by providing a standard guideline to leverage reusable components to attain the desired & reliable results rapidly. There are various factors for which an automation framework required [1].

- The Ease of use (Reusability of test script) to maintain the test script.
- To maintain reliability and consistency follow specific guidelines.
- To maintain scalability with changing requirement.
- To improve the efficiency of testing activities
- To provide easily understandable reports of automated results.

2.2. How do we select best appropriate test automation tool?

Selecting best appropriate tool for automation is one of the basic tasks for organization. The impact of selecting best automation tool may greatly affect the automation process which may provide sophistication in automation. While selecting tools for automation, certain conditions must keep in mind which makes sense for automation, like how frequently the repetitive test may occur, how often the regression testing is required etc. There are certain criteria to evaluate and select the automation tools [2] as mentioned below.

- Whether the tools fulfill the testing requirements. Is it compatible with project environment & technology that support objects in application?
- Which testing level and types does the tools support like (system, functional, performance) testing. Select tools which support maximum testing type with the ability to automate complex requirements.
- Whether the tools support all major browsers/OS or portable devices in which the application runs?
- Whether the estimated budget for purchasing automation tool affordable for organization?
- Whether the organization has skilled resources that can understand the language the tools support?
- Does the tool have powerful mechanism for graphical reporting?

While selecting a tool comprehensive understanding of its integration, reporting and ease of use with other automation tool must be executed. The standard automation testing framework should be application & tool independent & loosely coupled, it means that if the organization decides to migrate from one automation tool to other, it need not required to create a new test automation framework from scratch.

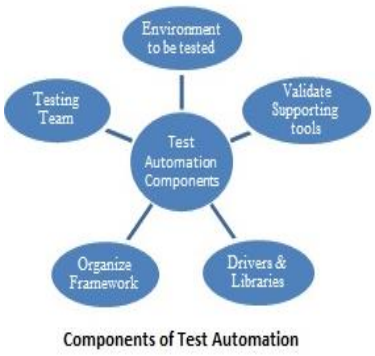


Figure 1: Components of Test Automation Framework

2.3. Components of test automation framework [3]

1) Environment to be tested: While designing a test automation framework all the test related tools, equipment, scripts and procedures must be verified and validated to ensure the system is stable to perform the desired task. This would be the first step towards the environment setup for the test automation. Until the test environment is not supposed to work accordingly the test automation may not be cost effective. Thus all the coordinating tools that support the environment must be stable and work together smoothly as a whole prior to start the test automation.

2) Validate supporting tools: While writing test scripts various supporting tools are required to organize the test. All such type of tools are properly investigated and independently monitored for expected contributory result as a unit, like test automation tool, bandwidth monitoring tools used for traffic management, other supporting tools such as requirement analysis, configuration management tools, test factory, defect tracking. A traffic monitoring and controlling activity of network is important to improve performance, efficiency and security of network resources. Network accommodates different kind of data traffic, identification of such traffic or congestion may lead the network administrator to optimize the network. Various types of network traffics are [4].

Table 1: Different Forms of Network Traffic

Traffic Type	Problem	Solution	Example
Busty Traffic	Consume high bandwidth starve applications	Fix constraint to limited bandwidth access	Download of FTP, Video Content
Latency Sensitive Traffic	Susceptible to compete for bandwidth results in poor response time	Fix min. & max. bandwidth range based on priority	Streaming application, VOIP, Video Conferencing
Interactive Traffic	Susceptible to compete for bandwidth results in poor response time	Prioritize over less essential traffic	SSL transactions, IM, Telnet Session
Non-Real Time Traffic	Consumes bandwidth during business hours	Schedule bandwidth during non-business hours	Email, batch processing applications

Configuration management tools are used to implement the changes (add, remove and update) occur during the test automation framework design to all associated system by implementing and enforcing to adopt such possible change. Test factory is an extensive use of test automation that minimizes the cost of testing and error rate along with supporting rapid scalability. Defect tracking is tracking the logged defects from beginning to end by testing, inspection or from customer feedback in a product with the aim to provide latest version of the product to fix the defect.

3) Drivers & Library: Different browsers require their own driver to communicate and control with the WebDriver. Some of the browsers are OS dependent but Firefox and Chrome are easily available across the platform. Download and unpack the drivers and add the location to you system PATH variable. The browser driver interacts with the respective browser by establishing a secure connection. While a test script is executed with WebDriver environment various activities performed in the background like:

- For every selenium command an HTTP request is generated to communicate with the browser driver.
- Drivers receive the HTTP request trough HTTP Server.
- HTTP Server decides the execution of instruction on the browser.
- As HTTP Server receives the execution status, it automates the script.

Browser automation library has been used to interact with other task required for automation. Library works as basic building block to establish a successful test environment to execute the test script by providing the required jar files, libraries, drivers and other supportive files.

Steps to add the Selenium Library to the project:

- Download & extract current release of Java Selenium library.
- Open the Eclipse IDE, go to the "Project→Properties"
- New dialog box open, Click on " Java Build Path" in the left pan, the to the register "Libraries"
- Click on "Add External JARs" in the right pan.
- Move to selenium library downloaded folder to select all .jar files by clicking "Open" button.
- Repeat this for all .jar files in the sub folder libs as well.

Once the library is configured with the project it can be reusable with the different tests, such utilities can perform different tasks for different test cases.

4) Organize Framework: Selenium is automated testing framework for web applications. To organize the framework following major steps can be taken into consideration [5].

- Select Programming Language: Select appropriate programming language as per the need of application under test and the developer/tester going to use the framework to write tests?
- Select Unit Test Framework: To design a framework a unit test framework is required. TestNG Compatibility with JAVA is highly recommended as it eliminates the limitation of earlier framework with the ability to write more flexible and more powerful tests. Other features TestNG are: annotation, grouping, sequencing and parameterization etc.
- Design the Framework Architecture: The architecture of framework is designed in such a way that it can be easily maintainable, scalable and sustainable in nature.
- Build the SeleniumCore component: SeleniumCore is designed to communicate between the web elements and browsers instances, it is used to create and destroy WebDriver objects, since it is not being taken into consideration by test writers.
- Build SeleniumTest Component: SeleniumTest components are all those test cases going use the classes provided by SeleniumCore. The design pattern used is Page Object Model (POM). It means for every page in the web application their corresponding a Page Object for them to organize the UI element in the pages.
- Select a Reporting Mechanism: Unless we get useful insights from the test result to take meaningful corrective action, test automation is of no use. A good report must include details about number of pass or failed test cases, passing rate, execution time and reason why test cases failed.

5) Testing Team: Successful test automation depends on efficient testing tools, standard testing process and skilled testing team to execute assigned responsibilities. Tools, process and testing team are the three key player of test automation. Software testing team is mainly engaged for test automation to utilize dedicated automation testing resources. The test team ensures that

the success of testing and automation effort based on the roles, duties and skill they provide.

- ✓ Manual Tester – Design and develop manual scenarios from requirement analysis.
- ✓ Automation Tester – Design and develop automation script.
- ✓ Lead Automator- Design and develop project automation architecture. Provide training and guidelines to testing team and collaborate to achieve the testing goal.
- ✓ Automation Environment Expert- Design, configure and update test environment and automation software.

Apart from the above role and responsibility testing team manages test platform, test tools, test cases and also provides tutorial to the users.

2.4. Characteristics of test automation framework

- If test automation framework is independent of Operating System, Programming language, tools and browsers, it would be considered as standard framework.
- The framework must be easily understandable without having programming skills.
- The framework must be easily maintainable and extendable.
- The Framework must execute test cases automatically, which includes test scenarios, test data, report generation, error handling and reporting final result.
- The framework must be able to execute tests without human intervention.
- Every test executed, must assign status of either Pass or Fail, failed test cases must include the failure data, the point of failure and other short description for failure.
- The framework should have meaningful logging and reporting structure.

2.5. Types of automation framework [6]

1) Linear Scripting Framework - This framework is also known as 'Record and Playback' framework as it generates the test scripts based on simulated record and playback method. It is not useful to run with multiple data set i.e. lack of reusability. This framework is suitable for small size applications. It works for individual test cases because it records the test script for particular scenarios.

2) Modular Testing Framework - In this framework, the complete application is divided into number of smaller modules and test script is generated individually. Finally the results of individual modules are compiled together by the master script to achieve the required target.

3) Keyword Driven Testing Framework – This framework is based on the keywords or actions for executing functions or methods. The test script is totally keyword/action dependent stored in tabular format in excel sheet, so it is also known as table-driven testing. The main class maintains the keywords and their corresponding action in the external table.

4) Data Driven Framework – Data Driven Test Automation framework is based on bifurcating the test data from test script. Test data is stored in separate files/format, test scripts connect and execute with different sets of test data. Due to having variant combination of test data set in separate table, this framework covers more test cases with flexibility in execution of test by changing input test data. This framework requires the programming skill to develop the test script.

5) Hybrid Driven Testing Framework – This framework is combination of two or more framework to leverage the strengths and benefits of each of the framework individually and to produce a better test environment overall.

6) Behavior Driven Development Testing Framework – The Purpose of using this framework is to share the process and tools and to communicate amongst stakeholders, with the aim to deliver quality product. BDD test cases are being written in simple language that can be easily understandable, which contains the details of how application behaves? BDD provides the platform for both technical and non-technical team to work together. A test case written in BDD follows the format 'Given-When-Then'. Given a certain scenarios, When an action takes place, Then this should be the outcome.

3. Related Works

Selenium WebDriver is powerful tool for test automation that directly interacts with web browsers, it follows the page object pattern to access the web page to reuse the test script to write the test cases. The page object pattern is used to reduce the coupling between web pages and test cases, which improves test suite maintainability [7]. UI locators are used to locate the web elements (Text box, button, checkbox, etc.) on the web page. Different identifiers are used to locate the web elements like (ID, link, XPath, link Test etc.), ID Locators have better result than XPath Locator, Combination of ID locator or XPath Locator with Link Text have improved the result but still not conclusive [7].

WebDriver tool can be used to write test cases easily and efficiently, it can helpful to develop and analyze the test cases using screenshot property, at the same time it is easily maintainable, reusable and extendable with the new requirement. Since the selenium WebDriver does not support any built in mechanism to generate test report, for that purpose TestNG is used as a testing framework. TestNG is compatible to perform unit, functional and integration testing. The author has generated a customized test report to analyze the failure test cases using screenshot [8].

WebDriver is cross platform testing framework that provides dynamic characteristics of web page, it means page reload is not required, if page elements gets changed. Selenium WebDriver execute faster as it directly interacts, run and control browsers using own control engine [9]. Selenium WebDriver is not equipped with reporting mechanism, TestNG generate report on behalf of selenium result. Such report contains execution time, method and detailed status of test script of running test suite "Passed or Failed" [10].

Table 2: Comparisons of Automation tools

Products & Parameters	Selenium	Katalon Studio	Test Complete	QTP/ UFT	Soap UI
Ease of Configuration & use	Advance skill required for configure & execution	Easy to configure & execute	Easy to configure. Training required for tool operation	Expertise require for configuration. Training required for tool operation	Easy to configure & execute
Licences required	Open Source & free to use, Free of Charge	No licence required. Not open source but free	Commercial tool, Licence required	Commercial, It is HP licenced product avail as single/Concurrent user	SoapUI is an open source. SoapUI Pro is the commercial.
Service support	Expert or group professional support	User friendly interface	Community Volunteer support	HP professional support	SoapUI pro Customer get support from forum or wen form
Operating System Support	Ms Windows, Apple OS X, Linux, Solaris	Ms Windows7/8/10, Apple OS X, Linux	MS Windows10/8.1/8/ 7/ vista, Server16/12/8	Only Windows Windows7/8/8.1/ XP/Vista/ 2003/2008	Ms Windows, Apple OS X, Linux
Scripting Language	Java, C#, Perl, PHP, Python, R, Ruby, JavaScript	Groovy, Java	JavaScript, C#, Python, VBScript, JScript	VB Script, Java Script, Window Shell Script	Groovy
Browser Support	Firefox, IE, Chrome, Safari, Opera, Phantoms, HtmlUnit, Android, iOS	IE9/10/11,Firefox, Chrome, Opera, Safari, MS Edge, Remote, Headless Browser	MS Edge,IE10/11, Chrome, Firefox	IE(V6-V11), Firefox(V3-V31),Chrome(V12-35),Safari6/6.1/7	SoapUI Default Browser
Device/Technology Support	Web apps	Web/API/Mobile/ Desktop apps	Desktop, Web & Mobile apps	Web/API/Mobile/ Desktop apps	API/Web apps
Testing Framework	RSpec (Ruby), Test::Unit (Ruby), Unittest (Python 2), JUnit 4 (Java), TestNG (Java), NUnit (C#)	Katalium Framework with TestNG	Application's UI	Build-in Feature for framework, like Key-word driven, Data Driven, Hybrid	Application's GUI
Type of test Performed	Functional Testing	Cross Browser Testing	Unit Testing, Functional Testing	Functional, Regression & Service Testing	Functional. Load & compliance Testing
Initially launched	2004	2015	1999	1998	2005

Comparisons of automation tools [11]

4. Framework Design for Automation of Cloud-Application Testing Using Selenium

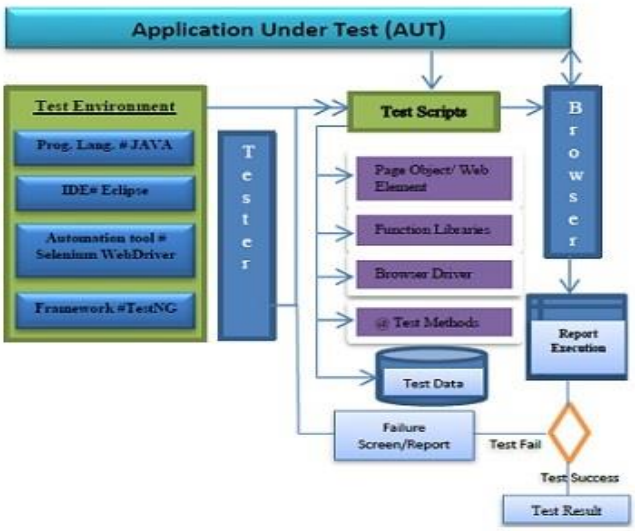


Figure 2: Architecture of Proposed Framework (FACTS)

5. Proposed Work

The framework (FACTS), we have designed is based on the consideration of test environment (includes language binding, IDE, automation tools and testing framework), types of Application under Test (AUT) with the compatibility of web browsers. The architecture of the proposed framework (Figure 2) includes following tools, functions, method and libraries etc.

- Programming Language # Java
- IDE #Eclipse
- Automation Tool # Selenium WebDriver
- Framework # TestNG
- Test Script
- Page Object /Web Element
- Function Libraries
- Browser Driver
- @Test Methods
- Test Data

The proposed framework may utilize the following tools, functions, method and libraries for designing automation framework:

- *Programming Language #Java*: Why Java is selected for this test automation framework? Java is a programming language which supports multiple operating systems. It is platform independent language, support Object Oriented features. Other important characteristics are simple, robust, Secure, Multithreading, distributed, portable to use and free of cost availability, open source and licence free.
- *IDE #Eclipse*: Eclipse is an Integrated Development Environment (IDE) tool most widely used to develop applications in Java and other programming languages

like C, C++, Perl, PHP, Python, Groovy, Ruby, R etc. Eclipse development environment is customizable as per the workspace and plug-in required. It is free to download and install.

- *Automation Tool #Selenium WebDriver*: Selenium is widely accepted open source test automation tool to test web-applications. Selenium WebDriver API is easy to explore and understand which support tester to easily read and maintain the test. Selenium WebDriver is web application automation framework that execute test across multiple browser, supports multiple operating system, support different programming language also. WebDriver executes faster due to their internal architecture to directly communicate with browser. When we execute a selenium command, an HTTP request created to communicate with particular browser driver, the browser driver then uses HTTP services to get HTTP requests, to establish the selenium command. Since WebDriver does not provide any build in facility to automatically generate Test result, we have incorporated TestNG as test reporting tool.
- *Framework #TestNG*: TestNG evolved as a testing framework to overcome the shortcomings of Junit and NUnit. It is an open source used for test automation framework. Where NG stand for Next Generation. According to some survey 76% tester uses Java programming along with TestNG framework. TestNG also provides an option to execute failed test cases only, it creates a file named testing-failed.xml for failed test cases in output folder. If the tester is interested to see only failed test cases they can execute this .xml file. There are various features due to which TestNG is in demand:
    - TestNG covers a broad range of test at different levels started from unit testing to system testing.
    - It supports annotations to create test cases easily.
    - It introduces 'test groups' to group test cases with prioritization, if required.
    - It supports parameterization and parallel test execution, it also execute multiple programs/classes using .xml.

The combination of test automation tools used for proposed framework has gained immense popularity among test automation tools as language, IDE and framework. The combination of the tools (i.e. **J**ava, **E**clipse, **S**elenium and **T**estNG) makes JEST for design and develop test automation framework.

- *Test Script*: Test script is set of instructions execute to test the system functions on the application under test for expected results. Test script mainly includes detailed descriptions of actions with required data to perform test. Test script follows the sequence to explore how to use the program and the order in which it is being utilized/executed, by pressing which button to carry out particular action in the program along with the stepwise expected results to observer the change in UI. Test script has a drawback with project which changes their requirement frequently, so the testers have to continuously update the script to fulfill the requirements. Since the test script is designed to test a particular task frequently, if bugs occur in the code the result would be same every time until the tester change the conditions.
  Page Object /WebElement: A Page object is an object oriented class through which AUT interface the web page. It plays a key role to identify the element on the web page, at the same time it reduces the code duplication by enhancing the mechanism of test maintenance. Whenever a test needs to communicate with UI of the page they use only methods of page object class. In case the UI of the page gets changed, only test script within the page object needs to be change. All possible user interactions can be used as a method on the class, ex. *clickLoginButton ();*
  setCredentials (user_name, Password);
  Selenium encapsulates elements on the page as an object of WebElement. WebDriver uses various techniques to identify the element on the webpage based on the its properties like ID, Name, Class, XPath, TagName, CSS Selectors, Link Text etc. WebDriver uses method like findElement() to find single web element and returns WebElement object. The method findElement () uses various locators like By.id (), By.name (), By.XPath () etc. to locate the element on the page by its properties.
  Ex. WebElementemail_id=driver.findElement (By.id ("User_Name")); // to get the web element for email address text field.
  Selenium WebDriver uses *findElement(By.Locator())* method to find the element on the page. Here findElement method uses locator/object as 'By' identifier.
- *Function Libraries*: One of the major purposes of designing a framework is its reusability, if any functionality is used frequently, such function can be written at one place and called multiple times for other test case. It is known as generic function. User defined functions are created based on their own requirement used with help of calling the appropriate package file. Function libraries are used to maintain the logic which is being used to write automation test scripts and called with their name. Methods are also known as functions.
- *Browser Driver*: Selenium WebDriver is web Automation tools that help to run the test across multiple browsers like FireFox, IE or Chrome. Corresponding browser driver is required to run the test in particular browser. Selenium launches corresponding browser as

per the script to execute test steps. To initialize WebDriver we can write code as :

WebDriver driver = new FireFoxDriver(); // where WebDriver works as interface and driver is reference variable for interface. Object can be associated with the instance of class FireFoxDriver//

When we execute a selenium command (driver.getTitle();) an HTTP request will be created to communicate with particular browser driver, the browser driver then uses HTTP service to get HTTP requests to establish the selenium command. All the built-in methods of corresponding drivers can only be accessible after configuring the driver.

- *@Test Methods*: Why do we use @ annotation before writing test method. Annotations in selenium are used to bind the associated method going to execute. Test annotation @ always used before every method, if this test annotation '@' is not prefixed before the test method, that method would not be considered as part of test code to be executed. @Test indicates that the annotated method is a part of test case serves as the main method [12]. Some other annotations used with TestNG are:

  @BeforeSuite: This method will execute once before all tests in this suite have run.

  @BeforeTest: This method will execute before any test method belonging to the classes inside the <test> tag is run.

  @BeforeClass: This method will execute only once before the first test method in the current class is invoked.

  @BeforeMethod: This method will execute before each test method.

  @AfterMethod: This method will execute after each test method.

  @AfterClass: This method will execute only once after all the test methods in the current class have executed.

  @AfterTest: This method will execute after all the test methods belonging to the classes inside the <test> tag have executed.

  @AfterSuite: This method will execute only once after all tests in this suite have executed.

- *Test Data*: Test data is actual input used to test the application. Test data would be used for positive or negative testing. Positive data is used to verify the functionality of application to meet the expected results. Negative data is used to test the ability of program to handle unusual, exceptional or unexpected input with its corresponding outcome [13].

6. Conclusion

In-depth planning, process and efforts are expected while design and develop a test automation framework. Selecting right automation tools plays a pivotal role to accomplish the automation task. Selenium WebDriver has been used to test web application end-to-end. WebDriver is flexible to integrate with variant development environment, tools (test, reporting and automation), OS, browsers and methods to design a test automation framework. Various types of test automation frameworks are available to design and develop test automation. The proposed framework (FACTS) is designed to work with Data Driven Framework as well as Hybrid Driven Testing Framework. A Standard automation tool is expected to execute its integration with other automation tools and reporting mechanism. Positive and negative test must be performed to verify the functionalities of application and to check the ability of program to handle unusual exceptions respectively.

References

[1] "A Guide To Automation Frameworks |Smartsheet" . Smartsheet, 2019, https://www.Swmartsheet.com/test-automation -frameworks-software.

[2] K. Kunal, and Mohammad Saad. "Test Automation Tool Selection Criteria And Comparison Matrix (A Complete Guide)". Softwaretestinghelp.Com, 2019, https://www. Softwaretestinghelp.com/automation-testing-tutorial-4/.

[3] Genius, Software. "Six Major Components of A Test Automation Framework - Software Testing Genius". Software Testing Genius, 2019, https://www.softwaretestinggenius. com/six-major-components-of-a-test-automation-framework/.

[4] "Network Traffic Management | Ipv6.Com". Ipv6.Com, 2019, https://www.ipv6.com/applications/network-traffic-management/.

[5] Logigear.Com, 2019, https://www.logigear.com/blog/test-automation/ building-a-selenium-framework-from-a-to-z/.

[6] "Types Of Test Automation Frameworks | Software Testing Material". Software Testing Material, 2019, https://www. softwaretestingmaterial.com/types-test-automation-frameworks/.

[7] Leotta, Maurizio et al. "Comparing The Maintainability Of Selenium Webdriver Test Suites Employing Different Locators: A Case Study". ACM, 2013, doi:http://dx.doi.org/10.1145/2489280.2489284.

[8] Gojare, Satish et al. "Analysis And Design Of Selenium Webdriver Automation Testing Framework". Procedia Computer Science, Vol 50, 2015, pp. 341-346. Elsevier BV, doi:10.1016/j.procs.2015.04.038.

[9] K, Revathi, and Janani V. "Selenium Test Automation Framework in On-Line Based Application -IJARSE". International Journal of Advance Research In Science And Engineering, Vol. No.4, no. Special Issue (02) ISSN-2319-8354(E), 2015, pp. 55-63.

[10] Singla, Sherry, and Harpreet Kaur. "Selenium Keyword Driven Automation Testing Framework ". International Journal Of Advanced Research In Computer Science And Software Engineering, Volume 4, Issue 6, 2014, pp. 125-129.

[11] Nurul, Md., and S.M.K. Quadri. "Software Testing Approach for Cloud Applications (STACA) – Methodology, Techniques & Tools." 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Jan. 2019, 10.1109/confluence.2019.8776915.

[12] "Selenium WebDriver Tutorial | TestNG For Test Case Management And Report Generation | Edureka". Edureka, 2019, https://www.edureka.co/ blog/selenium-webdriver-tutorial.

[13] "Test Data Generation: What Is, How To, Example, Tools". Guru99.Com,2019,https://www.guru99.com/ software -testing-test-data.htm

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 229-243 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Attacks classification and security mechanisms in Wireless Sensor Networks

Amine Kardi[1,*], Rachid Zagrouba[2]

[1]*Faculty of Mathematical, Physical and Natural Sciences of Tunis, University of Tunis El Manar, 2092 El Manar, Tunisia*

[2]*College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia*

ARTICLE INFO

Article history:
Received: 12 September, 2019
Accepted: 16 November, 2019
Online: 05 December, 2019

Keywords:
Wireless Sensor Networks
Routing attacks
Cryptography
Security
NS3 simulator

ABSTRACT

This paper proposes a new classification model distinguishing four classes of attacks in Wireless Sensor Networks (WSNs) namely: attacks based on the protocol stack, on the capability of the attacker, on the attack impacts and on the attack target. Then, it presents and classifies the most known attacks in WSNs based the proposed model. Simulations implemented under the NS3 simulator prove that the network lifetime can decrease by more than 45% in the presence of attacks. Afterwards, it discusses the main security methods and protocols of management and distribution of encryption keys used to ward off different types of attacks. Obtained results confirm that these security methods must be adapted to the specific characteristics of WSNs to achieve the intended objectives.

1. Introduction

With the widespread use of WSNs[1][2], composed of a small and low-powered communicating devices usually deployed in hostile environments, and due to their resources constraints, security factors become crucial and challenging with the appearance of a variety of attacks targeting these networks [3][4][5]. The main contribution of this paper is to provide a survey of these attacks and of the security techniques which will facilitate the design of WSNs for researchers and routing protocol programmers.

The rest of the paper is organized as follows: Security requirements for WSNs are discussed in Section 2. Section 3 introduces our proposed taxonomy and discusses each category of attacks. Section 4 presents the main attacks in WSNs and analyzes them using NS3 simulator. The basic security mechanisms in WSNs are presented in Section 5 then followed by the analysis and findings in section 6. Lastly; we conclude the paper and highlight our future work in section 7.

2. Security requirements for WSNs

Due to the lack of a wired communication medium, the limited resources of sensors and the properties of the deployment [6][7], WSNs, deployed mostly in hostile environments with mission-critical tasks, have several security challenges which are more complex than with other types of networks. As an ad hoc network, security requirements for WSNs include the standard security metrics known as CIAA (Confidentiality, Integrity, Authentication and Availability) in addition to the security requirements specific to this type of network which aim to protect the information and resources from attacks and misbehavior. In the following we discuss the security properties and requirements we would like to achieve in order to establish a reliable communication in WSNs and to provide secure services [8].

2.1. *Standard security requirements*

As with other types of networks, standard security requirements are needed in WSNs such as:

- **Confidentiality:** This is the most important issue in network security which guards data from eavesdroppers. It ensures that a given message remains hidden and cannot be used by anyone other than the desired receiver. It protects packets from undertaking traffic analysis, passive attackers and modification. The standard approach for achieving this requires use of encryption techniques [9].
- **Integrity:** Data integrity is needed to ensure the reliability of data parquets. In fact, sent packets must be protected from modification (adding, altering, deleting, damaging or losing) by malicious intermediate nodes or by wireless channel disturbance during the transmission [10].

*Amine Kardi, Tunis, (+216.96.27.90.33),Email : amine.kardi@fst.utm.tn

- **Authentication:** all applications require data authentication which ensures the reliability of the message by giving the ability of each communication host to identify its origin and to verify the other's identity to counter to packet injection or spoofing and to all malicious routing information. Critical characteristics of sensor nodes and the wireless nature of the communication medium in WSNs make extremely challenging to ensure authentication in these networks [11].
- **Availability:** Availability affects several sides in the sensor network. In fact, it determines whether a node can use the resources if needed, whether data and services are available for on-demand use and whether the network is always available to ensure communication even in the presence of malicious attacks. A loss of availability may have serious impacts and threat the entire network, e.g. it may open a backdoor for malicious invasion in some cases and sensed information may become useless or of lower value, but also providing availability can hurt the network lifetime especially with limited energy resources [12].

2.2. *Specific security requirements*

Node characteristics and deployment environments make WSNs vulnerable to special attacks and subsequently several security requirements specific to this type of networks are needed such as:

- **Authorization:** ensures that only authorized nodes can manipulate data (provide, update...) in the network and prevents unauthorized access to resources [13].
- **Nonrepudiation:** it prevents a node from denying sending a message it has previously sent through the network [14].
- **Data freshness:** implies that used data are recent and valuable and cannot be replayed by a malicious actor, under any circumstances, after abandonment to prevent, i.e. overloading the network by sending previously captured packets. Timestamps and time-related counters can be used to ensure data freshness [15].
- **Transparency:** after leaving the network a sensor node should not be able to replay old packets or to read any future messages in the same way as a joining node should not be able to reload or to read any previously transmitted message in order to protect the network from malicious injected nodes [16].
- **Self-Organization:** In the absence of an overall network management infrastructure, sensor nodes which have the ability of self- organization may encounter several hazards and risks threatening the safety and operation of the entire network [17].

  Time Synchronization: due to its limited energy resources, sensor nodes may be turned off for periods of time in the form of a smart sleep in order to conserve energy and, in some application, the end-to-end delay of some packets is of a great extent that is why collaborative sensor nodes require a synchronization system which must be ensured by a secure synchronization protocol to deal with attacks that attempt to affect proper synchronization between nodes [18].
- **Secure Localization:** In several applications, sensor networks are designed to locate faults. Consequently, the network must be able to accurately and automatically locate each sensor node that is why location information must be handled in a secure manner to avoid attacks taking advantage of the security weaknesses of this information [19].
- **Anonymity:** malicious actors do not have to decrypt the source node of a packet in order to secure nodes and the sensing area which can be affected by erroneous data [20].
- **Survivability:** Network services and functionalities must be maintained even with the low levels required in the case of failures, attacks and compromised or destroyed nodes [21].

According to these security requirements, it's obvious that attacks in WSNs can be of different types and can affect the entire network system such as nodes, packets, data, paths, routing etc. The following section presents our proposed taxonomy.

3. Taxonomy of attacks in WSNs

Because of its specific characteristics, WSNs are vulnerable to various types of attacks. These attacks are of different mechanisms, techniques and goals which makes necessary to find and adopt a well-defined taxonomy in order to facilitate design and development of WSNs for researchers and protocol programmers.

According to the security requirements mentioned above, and taking into account the mechanisms and parameters of attacks; we propose a new taxonomy which divides attacks in WSNs into four major classes namely: **attacks based on the protocol stack, based on the capability of the attacker, based on the attack impacts and based on the attack target** as shown in figure 1.

A detail overview of each security attack class is discussed in the rest of this section.

3.1. *Attacks based on protocol stack*

The protocol stack used by nodes in a wireless sensor network consists of the Physical Layer, Data Link Layer, Network Layer, Transport Layer, Application Layer, power management plane, mobility management plane and task management plane as shown in figure 2. It ensures the energy efficiency in the network, manages the use of the wireless medium with the routing task and facilitates synchronization and cooperative efforts between sensor nodes.

Each level on the protocol stack takes on a set of tasks and guarantees a set of services. Therefore, it will be targeted by several attacks.

- Physical Layer:

This layer ensures the data transmission services which encompass the selection of access channels, radiofrequency regulation and deflection, signal processing and data encryption to increase the communications reliability. The importance of this layer exposes it to a variety of attacks targeting data privacy, resource depletion, signal and communications interruption, and so on [22].

- Data Link Layer:

This layer is responsible of physical addressing, error detection and/or correction, data streams multiplexing, medium access and

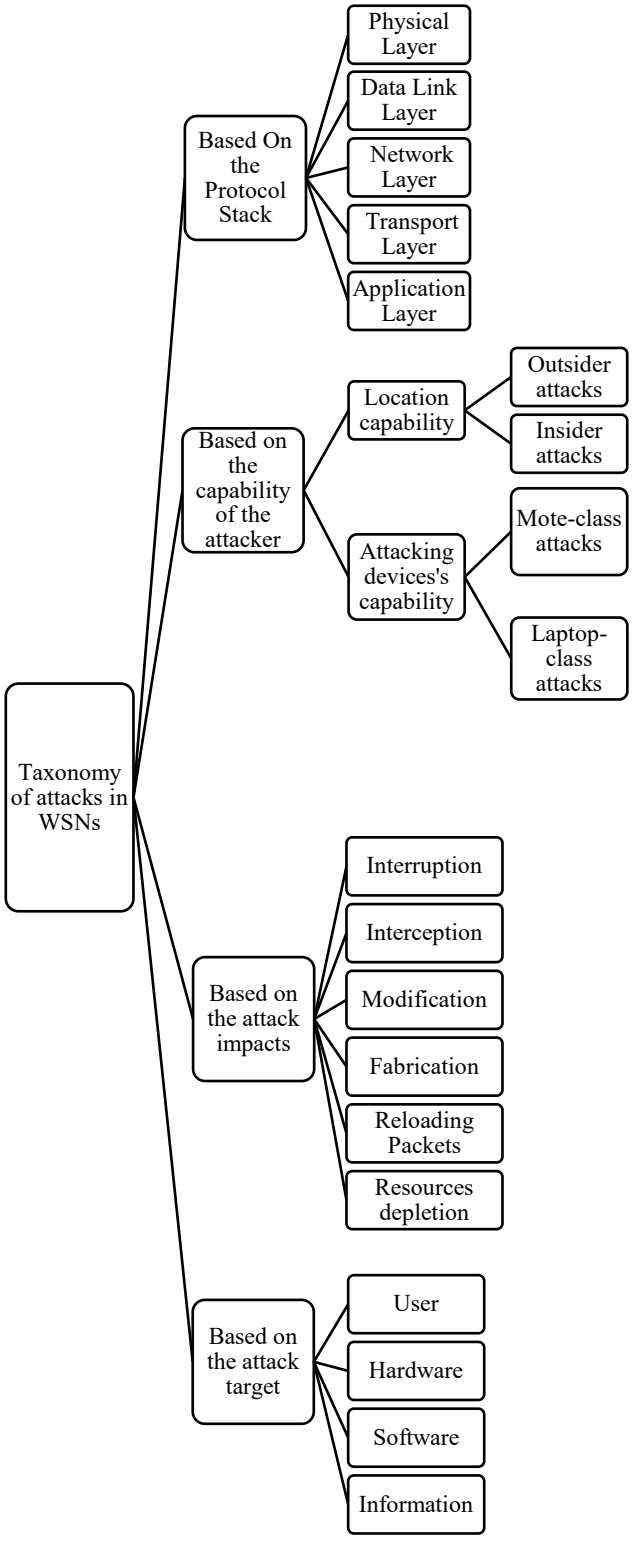


Figure 1: Security attacks in WSNs: A taxonomy

flow control as well as ensuring reliable point-to-point and point-to-multipoint connections in the network. The medium access control (MAC) protocol must support mobility and be able to handle collisions and to achieve energy efficiency in the network. The importance of this layer makes it vulnerable to several types of attacks sources of collisions, resource exhaustion, and network destruction [23].

- Network Layer:

This layer is in charge of routing the data supplied by its upper layer. It is very important task encompasses packets routing and forwarding, addresses assignment and energy management. These axial and fundamental tasks attract several attacks aimed at absorbing the network traffic, disrupting communication, exhausting resources, intercepting paths and injecting malicious packets [24].

- Transport Layer:

The transport layer ensures the management of end-to-end reliable delivery of packets connections and the establishment of end-to-end connection in addition to flow and congestion control. It is especially needed when the system is planned to communicate with external networks. Several attacks target this layer in order to reduce the network's availability, to degrade or even prevent data exchange and to waste energy [25].

- Application Layer:

This layer contains user applications which oversee the sensing tasks. It manipulates user data with a set of application protocols, and it is a target of attacks aiming at affecting the synchronization of communications and data confidentiality.

The power management plane and the mobility management plane are responsible, respectively, for the management of energy resources and the mobility of nodes in order to achieve maximum energy efficiency and stability of communications. The task management plane manages the sensing tasks and ensures synchronization between sensor nodes [26].

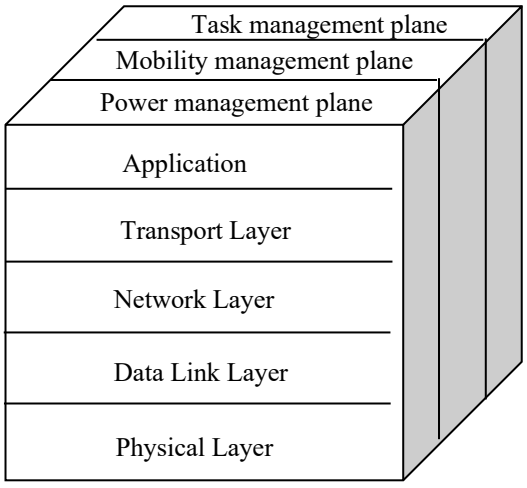


Figure 2: The protocol stack of sensor networks

3.2. *Attacks based on the capability of the attacker*

Based on the capability of the attacker, attacks in WSNs can be subdivided into two classes: Location capability and Attacking

device's capability. Each of these classes is further divided into two broad subclasses according to the relative capabilities of the attacker which are respectively: Outsider/ Insider attacks and Mote-class/ Laptop-class attacks. In the following we detail each of these subclasses.

- Outsider/ insider attacks

Based on the location capability of attackers, attacks in WSNs can be classified into Insider and Outsider attacks. We are talking about insider attacks when a legitimate nodes belonging to the network and which are actually part of it are compromised and participates in unintended and/or unauthorized ways in the attack contrary to outsider attacks which are launched only from outside the networks application environment and performed by external nodes which do not belong to the WSN. Insider attacks are more severe when compared with outside attacks since compromised nodes contains secret information and may not be detected due to the possesses privileged access rights they had. These attacks target the entire network and affect the resources, data, communications and even the existence of the network. Some papers refer to internal and external attacks [27].

- Mote-class/ laptop-class attacks

Based on the Attacking device's capability, attacks in WSNs can be classified into Mote-class and laptop-class attacks. In mote-class attacks, attackers use malicious nodes with similar techniques capabilities and hardware proprieties to the network nodes in order to execute malicious codes aiming to affect the network functionalities (data, routing, paths, energy management...). Contrariwise, attackers use more powerful devices with great transmission range, processing power, and energy supply than the network's sensor nodes in laptop-class attacks e.g., a laptop which results in launching more serious attacks and consequently lead to more serious damage to the target network [28].

3.3. *Attacks based on the attack impacts*

Based on the attack impacts, attacks in WSNs can be subdivided into six classes: Interruption, Interception, Modification, Fabrication, Reloading Packets and Resources depletion. In the following we detail each of these four classes.

- Interruption

The attacker aims, through some techniques, to interrupt the network's functionalities. These attacks are varied and can target nodes, sinks, paths, routing task, resources depletion etc. which threaten services availability. Even in the presence of an attack, the network must be capable of maintaining its functionalities without interruption [29].

- Interception

These attacks are used when the attacker intended to eavesdrop on the network information. This threatens message confidentiality, network services availability and may result in the depletion of resources [30].

- Modification

These attacks threaten data integrity. Attackers intend to eavesdrop to information and to tamper with it in various ways: Injection, removal, modification etc. in order to disrupt the network functionalities and threaten the accuracy of services and treatments [31].

- Fabrication

These attacks are manifested by the injection of erroneous data into the WSN. Attacker aims to threaten message authenticity and network availability through the depletion of the energy resources of the network especially with wireless communications which are very energy intensive [32].

- Reloading Packets

In some cases, the attacker seeks to reload and read previously transmitted packets from a compromised node or from buffers of active nodes which threatens data confidentiality and packets transparency. These packets can be injected back into the network, which threatens messages freshness and wastes energy [33].

- Resources depletion

In a WSN, resources depletion leads to the complete collapse of the network. Attackers aim to waste limited energy resources of nodes using various processing and transmission tasks [34].

3.4. *Attacks based on the attack target*

Attacks in a WSN can target different actors and components of the network which are users, hardware, software and information [35]. In the following we detail each of these targets.

- User

Various users of a WSN like humans, robots etc. can be targeted by malicious attacks aiming to have an authorized access to the network system through legitimate compromised identifiers. These attacks seek to take control of the network, to use it for personal tasks, to use information or to destroy it.

- Hardware

Attacks targeting hardware in WSNs tend to compromise nodes for malicious intent, waste limited energy resources or destroy equipment. A compromised node may contain secret information and may be used for unauthorized access to the network.

- Software

Attacks targeting software in WSNs aim to eavesdrop information, control the network and disrupt the services and functionalities.

- Information

Attacks targeting information in WSNs are various e.g. the attacker can falsify the data supposed to be sensed in the sensing area, in other cases, sensed data can be targeted by attacks, packets in transit, authentication parameters etc. Attackers aim at affecting the network functionalities, services availability and data confidentiality.

Each of these classes helps to analyze and study attacks in WSNs from different perspectives. We detail, in the following section, the most known attacks and countermeasures to counter their effects.

4. Attacks in WSNs

Generally, attacks under these four classes, can be either active or passive. A passive attack is an attack that aims to gets exchanged

data in the network without interrupting the communication to remains hidden and not be discovered. Passive attackers listen and collect data that can be used later to start other types of attacks. While active attacks involve some modifications in the data stream. In fact, it seeks to modify, fabricate, inject a false stream or intercept the data or the communication to disrupt the normal functionality of the network. In the active mode the attacker is more aggressive and aims to damage the entire network [36].

4.1. *The most known attacks:*

Among the most known attacks in WSNs we find:

- Jamming:

Jamming attack [37] takes advantages of the sensitivity of the wireless medium to interferences and results in a denial of service in the network. Attacker identifies the radio frequencies used by the targeted WSN and tries to disrupt or block communications by emitting signals (unnecessary information) in the same frequencies that inhibit communication (messages transmission and/or reception) between nodes. This interference may be temporary, intermittent or permanent and can affect all or part of the network depending on the radio range of the jamming source which can have the same characteristics as the network nodes as it can be more powerful. Jamming can be of different types: it can be constant if it targets and corrupts data packets in transit, deceptive when sending a constant data stream in the network, random when injected data streams are dispersed over time and reactive if the jamming source sends a jam signal following the detection of traffic. Jamming can target different layers: It can interfere with radio frequency signals at the physical layer, as it can take advantages of the medium access control protocol causing malicious collisions at the link layer as it can target the network layer by injecting malicious packets. Defense techniques [38][39][40] against Jamming attack involve spread-spectrum techniques for radio communication such as frequency hopping to switch between many frequency channels, authentication mechanisms to detect malicious and replayed packets, the use of secure medium access control protocols etc.

- Tampering or destruction [41]:

WSNs are highly vulnerable to physical attacks as they are often deployed in unprotected areas. A Tampering attack will be triggered after physically capturing the node and aims to retrieve cryptographic material as encryption keys and the program code stored within a node which can be used later to trigger other types of attacks such as modifying routing information, creating duplicate data packets, tampering routing services etc. or to manipulate the captured node by installing a new code causing an abnormal behavior of the compromised sensor, controlled by the attacker, in order to disrupt the network. Destruction attack consists of removing the sensor from the network by destroying it resulting in isolated areas and even the destruction of the entire network. Defense techniques against this attack are based on the principle of considering the situation in which sensor nodes are compromised such as tamper-proofing the physical package of nodes.

- Continuous Channel Access (Exhaustion):

Exhaustion [42] belongs to DoS attacks. It aims to exhaust the batteries of the nodes in order to reduce the network lifetime. It consists to disrupt the Media Access Control protocol, by continuously injecting unnecessary packets in the network and requesting or transmitting data over the channel entail the waste of node energy in unnecessary retransmissions. A possible solution to this attack is the Rate Limiting at the MAC admission control which allows ignoring corrupted packets and excessive malicious requests [43] [44].

- Collision:

Collision attack [45], similar to the continuous channel attack, can be caused by malicious nodes by transmitting packets in the network in order to block / delay the communication between nodes which results, in addition to the throughput, in a waste of energy and a loss of data. It is difficult to detect Collision attack in WSNs because it is a short time attack using malicious packets which are similar to legitimate packets. Researches, such as [46], proposed several Collision detection techniques such as error-correcting codes.

- Unfairness:

Unfairness attack [47] is a partial or a weak form of DOS attack that can result in marginal performance degradation. Based on other attacks such as collision and exhaustion, collision attack decreases significantly the utility and efficiency of services. In fact, attackers request continuously to access to the channel which results in undermining communication and limiting channel capacity. One of the countermeasures to such an attack is time-division multiplexing [48] allowing each node to transmit only in a specific time slot.

- Interrogation [49]:

To soften the problem of frame collisions introduced by the hidden node problem, several MAC protocols use the RTS/CTS handshake [50]. Attackers take advantages of this synchronization mechanism by repeatedly sending RTS packets leading to CTS responses from neighboring nodes. To counter this type of attack, nodes can use several mechanisms such as anti-replay protection and link layer authentication [51].

- Sybil Attack [52]:

As shown in figure 3 Sybil Attacks allow malicious nodes to have multiple identities by using the identities of the nodes targeted by the attack in order to participate in distributed algorithms such as election to take advantage of legitimate nodes to endorse the creation of several routes passing through the malicious node. This attack is located between the link and the network layers and it aims to degrade the integrity of data, the security level of and the use of resources. Among these attacks we can found Data Aggregation attacks aiming to falsify the aggregated message and Voting attacks affecting routing path and node selection. Several researches, such as [53] [54], sought to counter this attack using different mechanism such as public key cryptography and digital signatures.

- Sinkhole:

Sinkhole attack [55] belongs to DoS attacks in WSNs. As shown in figure 4, the attacker (malicious node) must appear to other nodes as being very attractive by presenting optimal routes in order to attract packets as much as possible to control most of

the data circulating in the network. Consequently, malicious node acts as a base station by attracting data packets and preventing them from continuing their path creating a kind of well or black hole in the network. The attacker, having a strategic routing location in the network, manipulates packets as desired, causing the suspension of the network routing service and the depletion of critical resources of nodes. Geo-routing protocols, using localized information and routing traffic through the physical location of nodes, are resistant to sinkhole attacks.

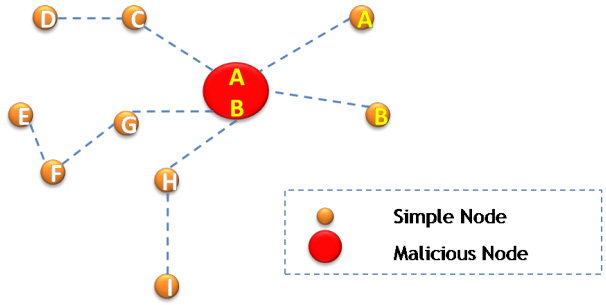


Figure 3: Sybil Attack

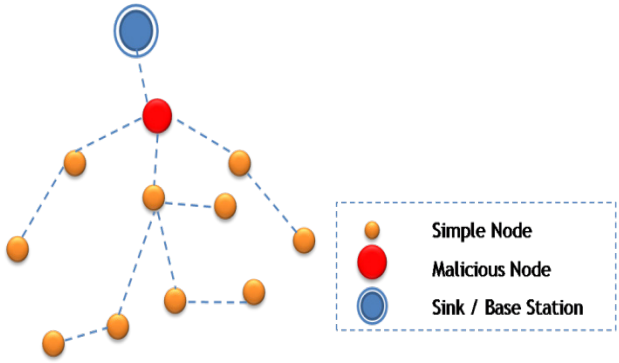


Figure 4: Sinkhole attack

- Hello Flood [56]:

The presence of laptop-class attackers and the limited transmission range of sensors led to the appearance of the Hello flood attacks. This attack takes advantages of the powerful signal of the attacking device to broadcast the information of an optimal route to all the nodes which will update their local tables. When a targeted node wants to communicate, it will not be able to use this route because the attacker, which is the imaginary neighbor, is out of range which results in the waste of energy from the sensors which leads to a malfunction of the network. Security mechanisms such as cryptography and source authentication, used to verify the bidirectionality of a path before taking action received over it, are used to counter this type of attacks.

- Node Capture:

In most applications, sensor nodes are highly vulnerable to physical attacks as they are often deployed in open areas easily accessible to attackers which raise the possibility of node capture to be used in tampering attacks. Researches such as [57] show that even a single node capture led to take over the entire network by attackers which makes the resilience against node capture attacks one of the most challenging issues in WSNs [58].

- Selective Forwarding:

Routing protocols assume that nodes will faithfully relay the packets that pass through them. We talk about Selective Forwarding attack when attacker may create malicious nodes and can violate this rule by removing all or some of these incoming packets randomly (neglectful node) or by giving high priority to its own messages (greedy node). Multipath routing, braided paths and random selection of paths to destination are a possible defense against this type of attack [59] [60].

- Black Hole Attack:

This attack can be considered as a form of selective forwarding attack dropping all incoming packets, it consists of inserting a new node or compromising a network node which falsifies the routing information to broadcast itself as the shortest path to oblige a maximum of neighbors to modify their routing tables in order to force the data to pass through it. This information will be destroyed which makes the malicious node as a well or black hole in the network. When attacker is placed on a strategic routing location in the network such as sinks or base stations the attack cause a network partition, packets loss, resource exhaustion and even the suspension of the routing services of the entire network. Black hole attack often targets hierarchical architectures and more specifically aggregator nodes. Multiple paths routing presents a defense against black hole attack [61] [62].

- Wormhole Attacks [63]:

Also known as tunneling, it is one of the most severe attacks. As shown in Figure 5, it requires at least two malicious nodes linked by a powerful radio link or by a wire link. In a wormhole attack, information received by a malicious node in one side of the network are encapsulated and relayed to be reintroduced by another malicious node on the other side of the network, revealing that the message originates from a close node. Encapsulation can be done in two ways: "Multi-hop encapsulation" aiming to hide intermediate nodes located between the two attackers to make the paths through the malicious node appear shorter which facilitates the creation of sinkholes with protocols using the number of hops as main metric of paths selection and "direct communication" in which paths going through the attackers are faster because they are composed of a single hop and can be used against protocols using the first discovered path and those based on paths latency. This type of attacks can considerably disrupt a location system by introducing erroneous reference points. It can be further divided into three classes: Half Open Wormhole, Closed Wormhole, and Open Wormhole. Several mechanisms for detecting and defending against wormhole attacks are presented such as WRHT [64].

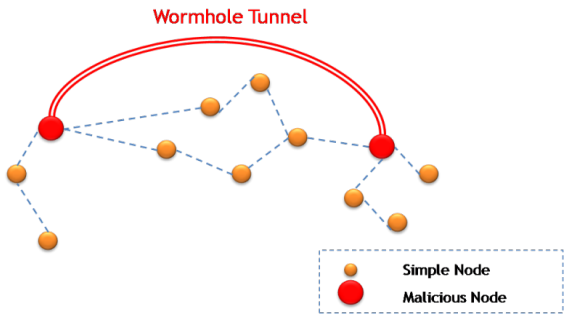


Figure 5: Wormhole attack

- Spoofed, Altered, or Replayed Routing Information [65]:

routing information in transit are targeted by several attacks, they can be spoofed, altered, or replayed. Malicious nodes can attract or repel the network traffic, inject false and misleading packets, overflow routing tables and create routing loops which result in disrupting the network traffic, partitioning the communication system and even destruction on the entire network. Several researches, such as [66], aim to counter these attacks using counters, timestamps, encryption and various authentication techniques.

- Misdirection [67]:

This active attack consists in misdirecting the data packets. In fact, instead of sending the packets in a correct direction, malicious nodes redirect them in a wrong direction through which the right destination is unreachable. Targeted node, flooded without any useful information, suffers from a waste of resources. Some security mechanisms [68] such as smart sleep [69] avoids this attack.

- Homing [70]:

This attack aims to locate critical nodes in the network providing essential services such as cluster heads and key managers by monitoring and analyzing the network traffic and nodes activities in order to eavesdrop their activities, take control of the network and extract sensitive data. To prevent homing attacks, protocols designers use different types of cryptographic schemes, algorithms and hide management messages [71] [72].

- Flooding [73]:

This attack aims to provoke a denial of service in order to decreases network lifetime. Indeed, one or several malicious nodes propagate many connection requests and carry out a regular massive sending to a strong emission power until exhausting the resources required the connection or reaching a maximum limit in order to saturate the network and prevent legitimate nodes from establishing communications which exhaust the resources (memory and energy) and reduce availability. One proposed solution to this attack is to demonstrate the commitment to this connection or to put a limit on the number of connections from a particular node [74] [75].

- De-synchronization Attacks [76]:

It is a part of the resource depletion attacks. The attacker causes missed frames by distorting, changing or increasing the sequence number of exchanged packets between end points. By receiving the modified packets, the destination deduces that the packets have been lost and requests the forwarding of these packets to the senders which disrupts the synchronization of communications and led to a considerable waste of energy of legitimate sensors by attempting to recover from errors which never really existed. Packets authentication is one of the main solutions to counter this attack [77] [78].

- Overwhelm attack [79]:

It is a part of QoS attacks. It consumes network bandwidth causing resource depletion by forwarding large amounts of data packets due to an exaggerated stimulation of sensors. A good adjustment of sensors and stimulation parameters, rate-limiting and efficient data-aggregation algorithms help to counter this attack [80].

- Path-based DOS attack [81]:

In this attack malicious nodes inject spurious or replayed packets in the network which waste energy resources and bandwidth on the path to the base station. Hence, legitimate nodes are prevented from sending packets to the base station. Authentication techniques and anti-replay protection counter this attack [82].

- Deluge (reprogram) attack [83]:

This attack uses one or more compromised nodes that the attacker reprograms them and then reinsert them on the network as malicious nodes working in his service. Deluge attack aims to disrupt the network and the application by causing an abnormal behavior of nodes. Indeed, these nodes run malicious codes in order to steal secret or sensitive information, attempt to disrupt the normal operation of the entire system and take control of large portions of a network. Several researches such as [84], introduce various techniques to counter this type of attacks.

Table 1 classifies the previously detailed routing attacks on WSNs according to our taxonomy proposed in the previous section.

4.2. *Simulation results*

4.2.1. Network model

To show the effect of these attacks on WSNs, we used the NS3 simulator to simulate the effect of Jamming and Hello Flood attacks on a WSN using LEACH protocol [85].

As indicated in Table2, the adopted network model consists of 100 nodes randomly deployed in 100m * 100m area with unlimited power sink centered in the field as shown in Figure 6.

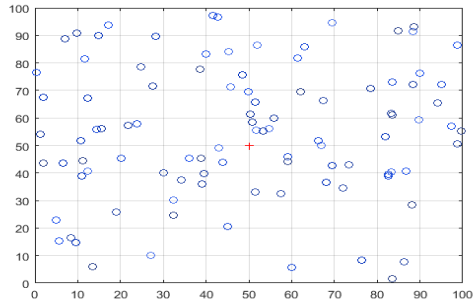


Figure 7: Nodes distribution

4.2.2. Results and discussion

In this section, the following metrics are used to evaluate the effect of Jamming and Hello Flood attacks:

- Network lifetime (Alive nodes Versus Rounds + Dead nodes Versus Rounds)

- Stability /instability periods

- Remaining energy Versus Rounds

We start with the evaluation of the network lifetime.

Table1: Comparison table of routing attacks in WSNs

Types / Attacks	Based On Protocol Stack					Based on the capability of the attacker				Based on the attack impact						Based on the attack target				Passive /active	
	PL	DLL	NL	TL	AL	LC		DC		IRP	ITC	M	F	RP	RD	U	H	S	I	P	A
						O	I	M	L												
Jamming	*	*	*			*	*	*	*	*					*		*				*
Tampering or destruction	*						*	*		*	*						*	*	*		*
Continuous Channel Access (Exhaustion)		*				*	*	*	*						*		*				*
Collision		*				*	*	*	*	*			*		*		*				*
Unfairness		*				*	*	*	*	*			*		*		*				*
Interrogation		*					*	*	*	*					*		*				*
Sybil Attack		*	*				*	*		*	*	*	*		*		*		*		*
Sinkhole			*				*	*		*	*	*	*	*	*		*		*	*	*
Hello Flood			*			*			*	*	*		*		*		*		*		*
Node Capture			*				*	*		*	*	*	*	*	*		*	*	*	*	*
Selective Forwarding			*				*	*		*	*	*	*		*		*		*		*
Black Hole Attack			*			*	*	*	*	*	*				*		*		*	*	
Wormhole Attacks			*				*		*	*	*	*			*		*		*		*
Spoofed, Altered, or Replayed Routing Information			*				*	*	*	*	*	*	*	*	*		*		*	*	*
Misdirection			*				*	*	*	*	*				*		*		*		*
Homing			*			*			*		*								*	*	
Flooding				*			*	*	*	*			*		*		*				*
De-synchronization Attacks				*			*	*	*	*		*			*		*		*		*
Overwhelm attack					*	*		*					*		*		*		*		*
Path-based DOS attack					*		*	*	*			*	*	*	*		*		*		*
Deluge (reprogram) attack					*		*	*		*	*	*	*	*	*		*	*	*		*

PL: Physical Layer
DLL: Data Link Layer
NL: Network Layer
TL: Transport Layer
AL: Application Layer
LC: Location Capability
O: Outsider
I: Insider
DC: Device Capability
M: Mote Class
L: Laptop Class
IRP: Interruption
ITC: Interception
M: Modification
F: Fabrication
RP: Reloading Packets
RD: Resources Depletion
U: Users
H: Hardware
S: Software
I: Information
P: Passive
A: Active

Table 2: Network parameters

Parameters	Values
Network size	100m * 100m
Number of nodes	100
Sink location	50,50
Packet size	4000 bits
Initial energy	0.5 J
Energy Dissipation (Efs)	10 pJ/bit/m²
Energy for transmission (ETx)	50 nJ
Energy for reception (ERx)	50 nJ
Data Aggregation	5 nJ/bit/report
Number of rounds	2000

- Network lifetime

The network lifetime is one of the main metrics to be considered to evaluate the efficiency of the network. It is the period between the beginning of the network and the death of last node. The two complementary Figures 8.a) and 8.b) respectively show the evolution of the number of alive and dead nodes in the network using LEACH protocol with and without the presence of attacks.

As shown in Figure 8, the network lifetime using LEACH protocol without undergoing attacks is about 1560 rounds, while it does not exceed 1080 rounds in the presence of Hello Flood attack and 820 rounds with Jamming attack. Using a powerful signal, Hello flood attack broadcast the information of an optimal imaginary route according to which the sensors update their local

tables. This technique led to a waste of energy and to a malfunction of the network. As explained earlier in this survey, deceptive Jamming attack, used in this simulation, consist in sending a constant data stream in the network which leads to overloading the antennas of the sensors and therefore to a great loss of energy and the disappearance of the network in a short time.

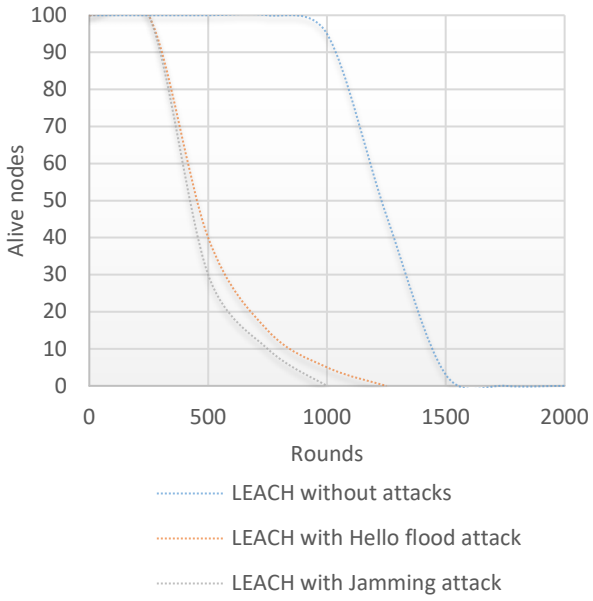


Figure 8-a: Alive nodes VS. rounds

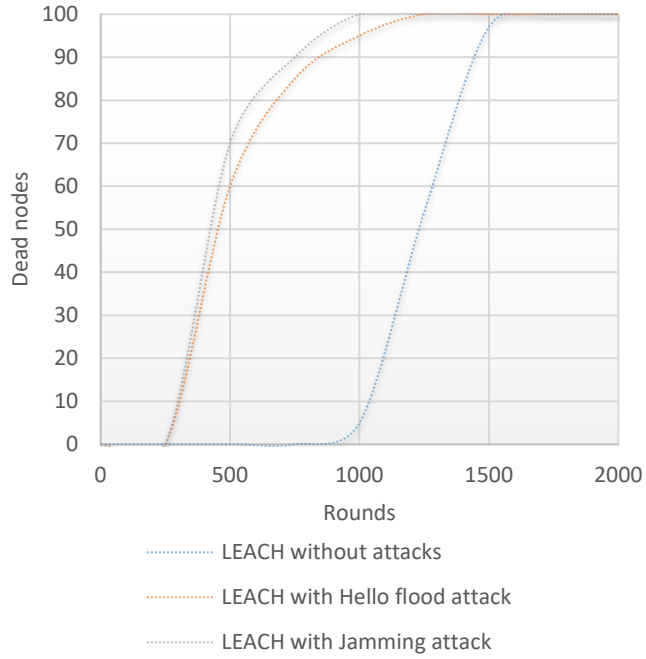


Figure 8-b: Dead nodes VS. rounds

Figure 8: Network lifetime

- Stability / instability period

A stability period triggers with the beginning of the network until the death of the first node which declares the beginning of an instability period. This instability period extends until the death of all nodes in the network. Figure 9 shows the stability and instability periods of the network using LEACH protocol with and without the presence of attacks.

As shown in Figure 9, the stability period almost reached 790rounds without attacks, but it does not exceed 320 rounds in the presence of Hello flood attack and 255 rounds with Jamming attack. In the same way, the instability period is about 773 rounds in the absence of attacks but does not surpass 760 rounds with Hello flood attack and 565 rounds in the presence of Jamming attack. This is obviously explained by the effect of the attacks on the consumption of energy.

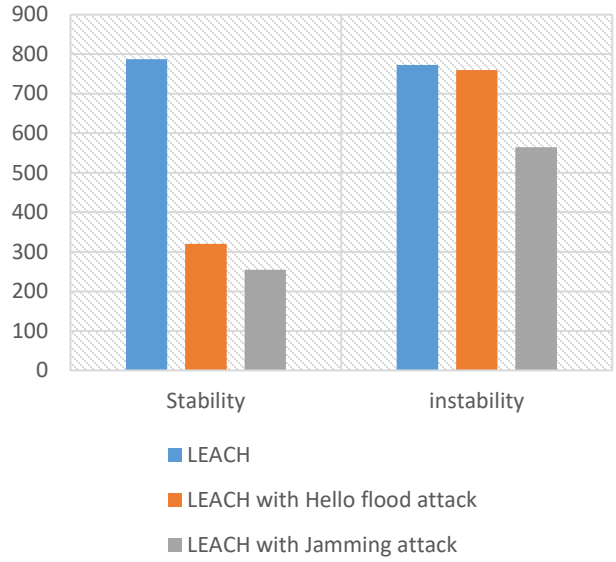


Figure 9: Stability and instability periods

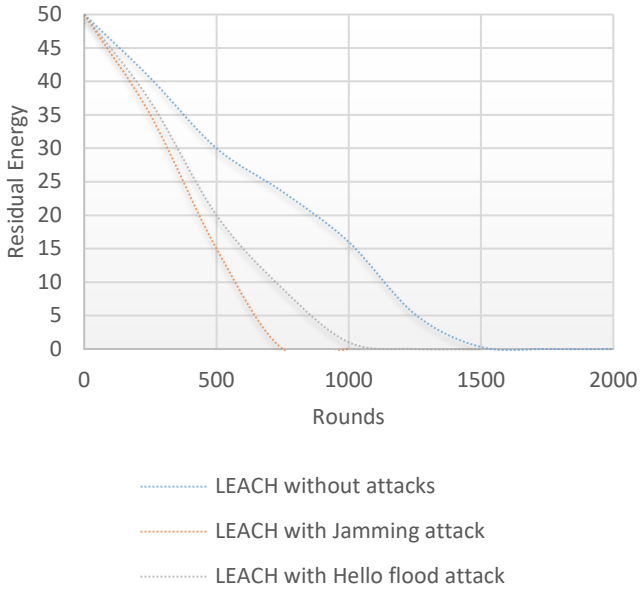


Figure 10: Residual Energy VS. Time

- Remaining energy

The residual energy, shown in Figure 10, results from the evolution of the number of nodes in the network. The curves show

the acute effect of the attacks on the residual energy of the network where half of the energy is consumed before reaching 500 rounds in the presence of attacks. The evolution of energy in this way leads to the formation of several isolated areas and subsequently affects the efficiency of the network.

In the following section, we present the main security mechanisms used to counter these attacks.

5. Basic security mechanisms in WSNs

In the absence of strong security architecture, WSNs are targets of several attacks that we presented previously in this paper. Most of these attacks take advantage of cryptographic system failures to derive the initial data (clear) or to find the used keys. We are talking here about cryptanalysis. In this section, we talk about cryptography and the cryptographic systems used to secure the WSNs [86] [87].

Indeed, the cryptanalysis targets the encryption keys which are based on cryptographic systems. These systems, that can be either symmetric, asymmetric or hybrid, take care of the management and distribution of the encryption keys [88] [89]. In the following, we detail each of these classes.

5.1. *Symmetric Cryptography in WSNs*

Symmetric cryptography, shown in Figure 11, is the oldest form of encryption where the same key is used between two communicating nodes to encrypt and decrypt the data using a symmetric encryption algorithm. In modern security techniques, even symmetric encryption algorithms can be public. The major disadvantage of this solution is that the pre-loaded key in the nodes could lead to compromise the entire network through compromised nodes. One of the proposed solutions to overcome this limit is to establish symmetric encryption schemes based on pair-wise keys rather than a single global key [90] [91].

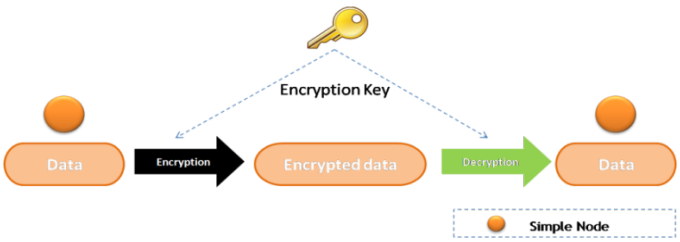


Figure 11: Symmetric cryptography

5.2. *Asymmetric Cryptography in WSNs*

Contrary to symmetric cryptography, asymmetric cryptography, also known as public-key cryptography, too heavy used in WSNs, uses a pair of keys instead of one (public and private). Asymmetric keys are generated in pairs. The message encrypted by the public key will be decrypted by the private key. Indeed, the public key of a node B, published between two communicating nodes is used by the node A to encrypt a message that will be decrypted by the node B through its secret private key as shown in Figure 12 and vice versa. This encryption mechanism allows not only the encryption of information but also the guarantee the authentication of the nodes by using the signatures and digital certificates. Slowness and resources consumption are the major drawbacks of this encryption mechanism [92].



Figure 12: Asymmetric cryptography

5.3. *Hybrid Cryptography in WSNs*

Hybrid cryptography is a cryptographic system that combines and benefits from the two cryptographic systems: asymmetric and symmetric. Generally, asymmetric encryption, greedy in time and resources, is used only to exchange a secret key in order to use symmetric encryption afterwards.

5.4. *Elliptic curve cryptography in WSNs*

Elliptic Curve Cryptography (ECC) is becoming increasingly popular as a security solution for WSNs. This technique is an approach of public-key cryptography, it is based on elliptic curve theory to create faster, smaller and more efficient encryption keys with quite less key size and very low computational overhead. In fact, the ECC generates keys through the properties of the elliptic curve equation and the discrete logarithm of an elliptic curve must be calculated for the decryption process which is much more difficult than factoring.

5.5. *Comparison of symmetric and asymmetric cryptographic systems*

Encryption algorithms can be evaluated in different parameters. Table 3, provided by National Institute of Standards and Technology NIST compares the encryption key size for RSA, ECC (asymmetric) and AES (symmetric) shows the development of key size compared to the security of an 80-bit symmetric key [93].

Table 3: Comparison of key length (bits) of the well-known symmetric and asymmetric techniques [NIST]

Symmetric (AES)	RSA	ECC
80	1024	160
112	2048(x 2)	224(x 1.4)
128	3072(x 3)	256 (x 1.6)
192	7680(x 7.5)	384(x 2.4)
256	15360(x 15)	521(x 3.2)

Researchers agree that symmetric methods are faster and require small keys e.g. a key size of 3072 bits with RSA, which is equivalent to 256 bits with ECC, offers the same level of security as a key of only 128 bits with AES. However, for a network of N nodes, N-1 symmetric keys must be stored in each node, which exceeds the memory capacity of sensors in networks that can contain thousands of nodes as well as the key distribution problem.

Figure 13, developed from experiments done using MATLAB R2015a [92] compares the energy consumed by RSA-1024 and ECC-160 for signatures generation and verification, and the energy cost of key exchanges excluding authentication and certificate verification. It shows that RSA is characterized by a very expensive signature operation and a small verification cost contrary to ECDSA signatures which are significantly cheaper than RSA. In the same way, key exchanging costs are significantly cheaper with ECDSA than with RSA.

As a result, the ECC has emerged as it offers keys that can be much smaller than those required by the RSA algorithm, and at the same time of size close to symmetrical solutions. These characteristics make this technique more suitable for WSNs and IoT applications but not optimal in its general form [94] [95].

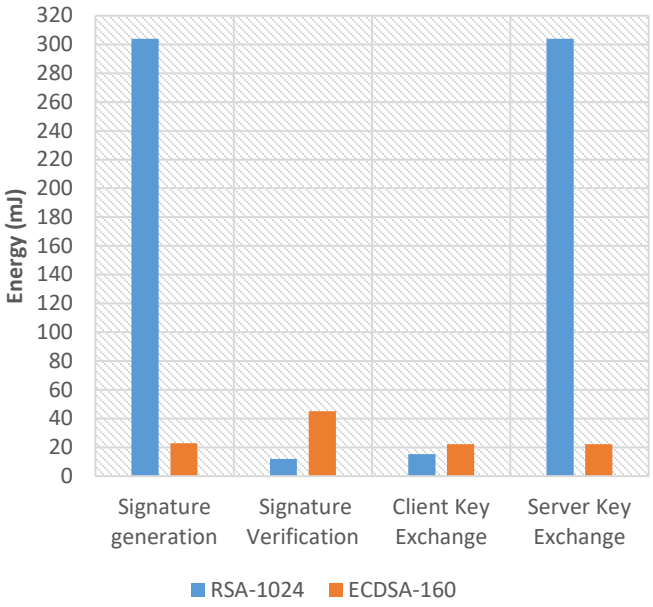


Figure 13: Energy cost of digital signature and key exchange using RSA-1024 and ECDSA-160

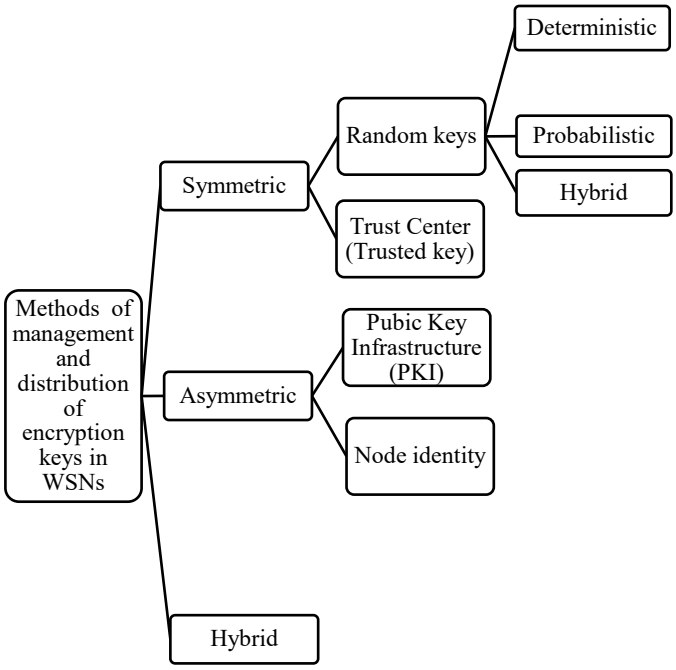


Figure 14: Methods of management and distribution of encryption keys in WSNs

5.6. *Management of encryption keys in WSNs*

Key management in WSNs is one of the most difficult aspects when configuring a security cryptographic system whether in a symmetric, asymmetric or hybrid-based method. This task includes the generation, distribution, verification and storage of encryption keys so that each node will be equipped with a set of secret keys or private/public pair of keys according to the adopted system in a secure, private and safe manner.

Several management classifications and key distribution have been proposed in the literature based on various technical and functional criteria [96]. We propose in this work a new classification based on the three previously proposed encryption methods (symmetric, asymmetric and hybrid) as shown in Figure 14. In the following we detail each of these key distribution methods.

5.7. *Management of encryption key in symmetric patterns*

The establishment of a common key between nodes in a WSNs to securely communicate in a symmetric cryptographic-based model is as follows: Before the nodes are deployed, the memory of each node is initially equipped with several keys in a Key Pre-distribution phase then it is the routing protocol that deals with the common key management between the nodes after the deployment in a key discovery process which makes it possible to constitute secure paths between nodes. The pre-distribution of these keys can be random or through an intermediary trust element of the network.

5.7.1. Random keys

The random pre-distribution of keys can be guaranteed by probabilistic, deterministic or hybrid methods. In the following we detail each of these methods

5.7.1.1. Deterministic methods

Deterministic methods require a large storage capacity because several encryption keys deterministically generated are stored in the memory of each node. They allow each node to establish a unique key known as “pair-wise key” with any other node of the network to establish a secure communication. Among the proposed mechanisms we quote the deterministic method proposed by Liu et al. [97] which is based on virtual spaces (logical grids) containing the identifiers of the nodes. Thus, many groups of sensor nodes are formed based on the positions of the nodes. The pre-distribution of two encryption keys is thus necessary: “in-group pre-distribution” making it possible to establish a unique shared key (pair-wise key) between two nodes of the same group and “cross-group pre-distribution” to establish a unique shared key between two different groups.

5.7.1.2. Probabilistic methods

In probabilistic methods, each node in the network is equipped with several encryption keys randomly chosen by the network administrator before being deployed. Common keys are thus installed in the memories of the nodes with a certain probability. In [98], Jones et al. propose a secure probabilistic architecture based on a probabilistic key distribution method. Indeed, each sensor is loaded with a set of “m” encryption keys, randomly selected from a set of “k” keys with a probability of matching "p" between nodes. After deployment, the BS splits the network into independent sectors and secret keys are distributed by the BS to ensure secure communication.

5.7.1.3. *Hybrid methods*

Hybrid pre-distribution methods combined and benefits from the two previous methods, depending on the application to increase connectivity between nodes while ensuring secure communication.

5.7.2. *Trust Center*

In this key pre-distribution model, a Trust Center e.g. the BS is used either to store a copy of pre-loaded keys or to derive encryption keys from a trusted key pre-loaded into the nodes. In a Zigbee topology managed by the protocol stack known as ZigBee [99] based on the IEEE 802.15.4 standard, the nodes can play three roles: coordinator, routers, and sensors. The coordinator, placed in a particular position in the network, plays the role of a Trust Center (TC) and distributes keys that can be Link Key (LK), Network Key (NK) or Master Key (MK) shared by one node respectively with another node, with the entire network and for the establishment of the key LK.

5.8. *Management of encryption key in asymmetric patterns*

Another way to establish a common key between two nodes or a group of nodes in a WSN to secure communications is to use schemes based on asymmetric systems as already explained. In the following we detail the most know methods.

5.8.1. *Public Key Infrastructure*

Public Key Infrastructure (PKI) based specifically on cryptography, are a set of protocols and services aiming to secure communication in a network through various techniques such as authentication, digital certificates, digital signatures etc. The Micro Public Key Infrastructure (micro-PKI) method proposed by Munivel et al. [100] for WSNs is based on two keys: a public key loaded in each node before the deployment used to identify with the BS and a private key used by the BS to decrypt the data sent by the nodes. The PKI in this method ensures identification between the nodes as well as with the BS.

5.8.2. *Node identity*

These methods seek to provide the same cryptographic services as a PKI based solely on the node identity information in the creation and establishment of cryptographic keys shared between each pair of nodes in the network. These methods are known in cryptography as the Identity-Based Non-Interactive Key Distribution Scheme (ID-NIKDS) [101]. Among these methods we quote the method proposed by Oliveira et al. [102] where each node of the network has a unique identifier and a private key. A unique secret key shared with another node of the network is derived by knowing only its identifier.

5.9. *Management of encryption key in hybrid patterns*

Hybrid cryptography is a system that combines the symmetric and asymmetric cryptography. It benefits from their advantages to ensure a high level of security in particular situations as mentioned previously in this work. The Transport Layer Security (TLS) [103] combines the two cryptographic systems. Indeed, as asymmetric cryptographic operations are very expensive, it is only used in the key distribution then a symmetric algorithm is used for the bulk encryption of the data.

In the following section, we summarize and discuss the results obtained in this survey.

6. Analysis and Findings

The specific characteristics of WSNs, usually deployed in hostile environments make these networks targeted by various types of attacks which can target all components of the network such as nodes, routes, packets etc. Simulation results show that the network lifetime can decrease by more than 45% as in the presence of Jamming attack. These conditions lead us to discuss the existing security techniques which, despite their proven effectiveness with other types of networks, cannot be directly adapted to the limited resources of sensors and a new security mechanism taking advantage of existing solutions must be proposed.

7. Conclusion and future work

As an involving field, WSNs usually used in unattended environments and characterized by limited resources become vulnerable to several types of attacks targeting all network components such as nodes, packets and routing protocols etc. In this paper, we aim to facilitate the design and implementation of reliable and secure routing protocols for researchers. So, we summarize the security requirements for WSNs then we propose a new taxonomy distinguishing four possible categories of attacks namely: attacks based on the protocol stack, based on the capability of the attacker, based on the attack impacts and based on the attack target. Then, we classify and analyze the most known attacks based on the proposed model and using the NS3 simulator. Thereafter, we present and discuss the basic security methods and protocols of management and distribution of encryption keys in WSNs.

Security Mechanisms in WSNs, used to counter these attacks in order to keep the network running smoothly will be the objective of our future work aiming to propose a new security method which will be adapted to the specific properties of WSNs and able to counter the majority of these attacks.

References

[1] YADAV, Devendra, SINGAM, Jayanthu, DAS, Santos, et al. A Critical Review on Slope Monitoring Systems with a Vision of Unifying WSN and IoT. IET Wireless Sensor Systems, 2019.

[2] PRABHU, Boselin, BALAKUMAR, N., et ANTONY, A. Johnson. Wireless Sensor Network Based Smart Environment Applications. 2017.

[3] SHAHZAD, Furrakh, PASHA, Maruf, et AHMAD, Arslan. A Survey of Active Attacks on Wireless Sensor Networks and their Countermeasures. arXiv preprint arXiv:1702.07136, 2017.

[4] SENGAR, P. et BHARDWAJ, N. A Survey on Security and Various Attacks in Wireless Sensor Network. International Journal of Computer Sciences and Engineering, 2017, vol. 5, no 4, p. 78-84.

[5] MITTAL, Vikas, GUPTA, Sunil, et CHOUDHURY, Tanupriya. Comparative Analysis of Authentication and Access Control Protocols Against Malicious Attacks in Wireless Sensor Networks. In : Smart Computing and Informatics. Springer, Singapore, 2018. p. 255-262.

[6] BANA, Suman et BAGHLA, Silki. Wireless sensor network. International Journal of Engineering Science, 2016, vol. 1706.

[7] Kardi, A., Zagrouba, R., & Alqhtani, M. (2018, May). Hierarchical Routing Techniques in Wireless Sensor Networks. In International Symposium on

Web and Wireless Geographical Information Systems (pp. 77-84). Springer, Cham.
[8] TOMIĆ, Ivana et MCCANN, Julie A. A Survey of Potential Security Issues in Existing Wireless Sensor Network Protocols. IEEE Internet of Things Journal, 2017, vol. 4, no 6, p. 1910-1923.
[9] CUI, Jie, SHAO, Lili, ZHONG, Hong, et al. Data aggregation with end-to-end confidentiality and integrity for large-scale wireless sensor networks. Peer-to-Peer Networking and Applications, 2017, p. 1-16.
[10] ORACEVIC, Alma, DILEK, Selma, et OZDEMIR, Suat. Security in internet of things: A survey. In : Networks, Computers and Communications (ISNCC), 2017 International Symposium on. IEEE, 2017. p. 1-6.
[11] DAS, Ashok Kumar. A secure and effective biometric-based user authentication scheme for wireless sensor networks using smart card and fuzzy extractor. International Journal of Communication Systems, 2017, vol. 30, no 1.
[12] MENDEZ, Diego M., PAPAPANAGIOTOU, Ioannis, et YANG, Baijian. Internet of things: Survey on security and privacy. arXiv preprint arXiv:1707.01879, 2017.
[13] PRITCHARD, Sean W., HANCKE, Gerhard P., et ABU-MAHFOUZ, Adnan M. Security in Software-Defined Wireless Sensor Networks: Threats, Challenges and Potential Solutions. In : IEEE Int. Conf. of Ind. Informat., Emden, Germany. 2017.
[14] SHAHZAD, Furrakh, PASHA, Maruf, et AHMAD, Arslan. A Survey of Active Attacks on Wireless Sensor Networks and their Countermeasures. arXiv preprint arXiv:1702.07136, 2017.
[15] ZHU, Liehuang, ZHANG, Zijian, et XU, Chang. Secure data aggregation in wireless sensor networks. In : Secure and Privacy-Preserving Data Communication in Internet of Things. Springer, Singapore, 2017. p. 3-31.
[16] MATHEW, Prabha Susy, PILLAI, Anitha S., et PALADE, Vasile. Applications of IoT in Healthcare. In : Cognitive Computing for Big Data Systems Over IoT. Springer, Cham, 2018. p. 263-288.
[17] TOMIĆ, Ivana et MCCANN, Julie A. A Survey of Potential Security Issues in Existing Wireless Sensor Network Protocols. IEEE Internet of Things Journal, 2017, vol. 4, no 6, p. 1910-1923.
[18] ZHANG, Anran et BAI, Fengshan. An energy efficient and dynamic time synchronization protocol for wireless sensor networks. In : Seventh International Conference on Electronics and Information Engineering. International Society for Optics and Photonics, 2017. p. 103221X.
[19] LI, Peng, YU, Xiaotian, XU, He, et al. Research on secure localization model based on trust valuation in wireless sensor networks. Security and Communication Networks, 2017, vol. 2017.
[20] LI, Xiong, NIU, Jianwei, KUMARI, Saru, et al. A three-factor anonymous authentication scheme for wireless sensor networks in internet of things environments. Journal of Network and Computer Applications, 2018, vol. 103, p. 194-204.
[21] ALIBERTI, Giulio, DI PIETRO, Roberto, et GUARINO, Stefano. Epidemic data survivability in Unattended Wireless Sensor Networks: New models and results. Journal of Network and Computer Applications, 2017, vol. 99, p. 146-165.
[22] ZHANG, Junqing, DUONG, Trung Q., WOODS, Roger, et al.Securing Wireless Communications of the Internet of Things from the Physical Layer, An Overview. Entropy, 2017, vol. 19, no 8, p. 420.
[23] DENER, Murat et BAY, Omer Faruk. Practical Implementation of an Adaptive Detection-Defense Unit against Link Layer DoS Attacks for Wireless Sensor Networks. Security and Communication Networks, 2017, vol. 2017.
[24] SAINI, Rakesh Kumar, RITIKA, Ritika, et VIJAY, Sandip. Data Flow in Wireless Sensor Network Protocol Stack by using Bellman-Ford Routing Algorithm. Bulletin of Electrical Engineering and Informatics, 2017, vol. 6, no 1, p. 81-87.
[25] KİRAZ, Ayhan et ÇAKIROĞLU, Murat. ALORT: a transport layer protocol using adaptive loss recovery method for WSN. Sādhanā, 2017, vol. 42, no 7, p. 1091-1102.
[26] LI, Xiaomin, LI, Di, WAN, Jiafu, et al. A review of industrial wireless networks in the context of industry 4.0. Wireless networks, 2017, vol. 23, no 1, p. 23-41.
[27] CHOI, Jaewoo, BANG, Jihyun, KIM, LeeHyung, et al. Location-based key management strong against insider threats in wireless sensor networks. IEEE Systems Journal, 2017, vol. 11, no 2, p. 494-502.
[28] TYAGI, Akshat, KUSHWAH, Juhi, et BHALLA, Monica. Threats to security of Wireless Sensor Networks. In : Cloud Computing, Data Science & Engineering-Confluence, 2017 7th International Conference on. IEEE, 2017. p. 402-405.
[29] GRGIC, Kresimir, MENDELSKI, Vedran, et ZAGAR, Drago. Security framework for visual sensors and smart camera networks. In : Telecommunications (ConTEL), 2017 14th International Conference on. IEEE, 2017. p. 131-138.
[30] JAITLY, Sunakshi, MALHOTRA, Harshit, et BHUSHAN, Bharat. Security vulnerabilities and countermeasures against jamming attacks in Wireless Sensor Networks: A survey. In : Computer, Communications and Electronics (Comptelix), 2017 International Conference on. IEEE, 2017. p. 559-564.
[31] SHAHZAD, Furrakh, PASHA, Maruf, et AHMAD, Arslan. A Survey of Active Attacks on Wireless Sensor Networks and their Countermeasures. arXiv preprint arXiv:1702.07136, 2017.
[32] BHUSHAN, Bharat et SAHOO, Gadadhar. Recent Advances in Attacks, Technical Challenges, Vulnerabilities and Their Countermeasures in Wireless Sensor Networks. Wireless Personal Communications, 2017, p. 1-41.
[33] SENNIAPPAN, Vijayalakshmi et SUBRAMANIAN, Jayashree. Threshold-based Energy Efficient Clustering Protocol for Corrosion Risk Monitoring of Reinforced Concrete. International Journal of Computer Applications, 2017, vol. 157, no 6.
[34] RAMACHANDRAN, Shyamala et SHANMUGAM, Valli. Impact of DoS Attack in Software Defined Network for Virtual Network. Wireless Personal Communications, 2017, vol. 94, no 4, p. 2189-2202.
[35] DESNITSKY, Vasily et KOTENKO, Igor. Modeling and Analysis of IoT Energy Resource Exhaustion Attacks. In : International Symposium on Intelligent and Distributed Computing. Springer, Cham, 2017. p. 263-270.
[36] SHAHZAD, Furrakh, PASHA, Maruf, et AHMAD, Arslan. A Survey of Active Attacks on Wireless Sensor Networks and their Countermeasures. arXiv preprint arXiv:1702.07136, 2017.
[37] VADLAMANI, Satish, EKSIOGLU, Burak, MEDAL, Hugh, et al. Jamming attacks on wireless networks: A taxonomic survey. International Journal of Production Economics, 2016, vol. 172, p. 76-94.
[38] SHABANA, Kalsoom, FIDA, Nigar, KHAN, Fazlullah, et al. Security issues and attacks in Wireless Sensor Networks. International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE), 2016, vol. 5, no 7, p. pp: 81-87.
[39] ZHU, Jia, ZOU, Yulong, et ZHENG, Baoyu. Physical-Layer Security and Reliability Challenges for Industrial Wireless Sensor Networks. IEEE Access, 2017..
[40] BHAVATHANKAR, Prasenjit, SARKAR, Subhadeep, et MISRA, Sudip. Optimal decision rule-based ex-ante frequency hopping for jamming avoidance in wireless sensor networks. Computer Networks, 2017.
[41] DINKER, Aarti Gautam et SHARMA, Vidushi. Attacks and challenges in wireless sensor networks. In : Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on. IEEE, 2016. p. 3069-3074.
[42] DINKER, Aarti Gautam et SHARMA, Vidushi. Attacks and challenges in wireless sensor networks. In : Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on. IEEE, 2016. p. 3069-3074.
[43] DENER, Murat et BAY, Omer Faruk. Practical Implementation of an Adaptive Detection-Defense Unit against Link Layer DoS Attacks for Wireless Sensor Networks. Security and Communication Networks, 2017, vol. 2017.
[44] MISHRA, Alekha Kumar. Security Threats in Wireless Sensor Networks. Handbook of Research on Advanced Wireless Sensor Network Applications, Protocols, and Architectures, 2016, p. 307.
[45] JAN, Mian Ahmad et KHAN, Muhammad. Denial of Service Attacks and Their Countermeasures in WSN. IRACST–International Journal of Computer Networks and Wireless Communications (IJCNWC), 2013, vol. 3.
[46] REN, Ju, ZHANG, Yaoxue, ZHANG, Kuan, et al. Adaptive and channel-aware detection of selective forwarding attacks in wireless sensor networks. IEEE Transactions on Wireless Communications, 2016, vol. 15, no 5, p. 3718-3731.
[47] MOHAMMADI, Shahriar et JADIDOLESLAMY, Hossein. A comparison of link layer attacks on wireless sensor networks. arXiv preprint arXiv:1103.5589, 2011.
[48] NEWAZ, SH Shah, CUEVAS, Ángel, LEE, Gyu Myoung, et al. Improving energy saving in time-division multiplexing passive optical networks. IEEE Internet Computing, 2013, vol. 17, no 1, p. 23-31.
[49] SINGLA, Aashima et SACHDEVA, Ratika. Review on security issues and attacks in wireless sensor networks. International Journal of Advanced Research in Computer Science and Software Engineering, 2013, vol. 3, no 4.
[50] KAUR, Ramandeep et SHARMA, Reecha. An Analysis of RTS/CTS Mechanism for Data Transfer In Wireless Network: A Review. 2016.
[51] GHEORGHE, Laura, RUGHINIS, Razvan, DEACONESCU, Razvan, et al. Authentication and anti-replay security protocol for wireless sensor networks. In : Systems and Networks Communications (ICSNC), 2010 Fifth International Conference on. IEEE, 2010. p. 7-13.
[52] LI, Xun, HAN, Guangjie, QIAN, Aihua, et al. Detecting Sybil attack based on state information in Underwater Wireless Sensor Networks. In : Software, Telecommunications and Computer Networks (SoftCOM), 2013 21st International Conference on. IEEE, 2013. p. 1-5.

[53] JAN, Mian Ahmad, NANDA, Priyadarsi, HE, Xiangjian, et al. A Sybil Attack Detection Scheme for a Centralized Clustering-based Hierarchical Network. In : Trustcom/BigDataSE/ISPA, 2015 IEEE. IEEE, 2015. p. 318-325.
[54] JAN, Mian Ahmad, NANDA, Priyadarsi, HE, Xiangjian, et al. A sybil attack detection scheme for a forest wildfire monitoring application. Future Generation Computer Systems, 2016.
[55] SHAFIEI, Hosein, KHONSARI, Ahmad, DERAKHSHI, H., et al. Detection and mitigation of sinkhole attacks in wireless sensor networks. Journal of Computer and System Sciences, 2014, vol. 80, no 3, p. 644-653.
[56] SINGH, Virendra Pal, UKEY, Aishwarya S. Anand, et JAIN, Sweta. Signal strength based hello flood attack detection and prevention in wireless sensor networks. International Journal of Computer Applications, 2013, vol. 62, no 15.
[57] LIN, Chi et WU, Guowei. Enhancing the attacking efficiency of the node capture attack in WSN: a matrix approach. The Journal of Supercomputing, 2013, vol. 66, no 2, p. 989-1007.
[58] TAHIR, Ruhma et MCDONALD-MAIER, Klaus. Improving resilience against node capture attacks in wireless sensor networks using icmetrics. In : Emerging Security Technologies (EST), 2012 Third International Conference on. IEEE, 2012. p. 127-130.
[59] REN, Ju, ZHANG, Yaoxue, ZHANG, Kuan, et al. Adaptive and channel-aware detection of selective forwarding attacks in wireless sensor networks. IEEE Transactions on Wireless Communications, 2016, vol. 15, no 5, p. 3718-3731.
[60] SHAHZAD, Furrakh, PASHA, Maruf, et AHMAD, Arslan. A Survey of Active Attacks on Wireless Sensor Networks and their Countermeasures. arXiv preprint arXiv:1702.07136, 2017.
[61] BASKAR, Radhika, RAJA, PC Kishore, JOSEPH, Christeena, et al. Sinkhole Attack in Wireless Sensor Networks-Performance Analysis and Detection Methods. Indian Journal of Science and Technology, 2017, vol. 10, no 12.
[62] ALJUMAH, Abdullah et AHANGER, Tariq Ahamed. Futuristic Method to Detect and Prevent Blackhole Attack in Wireless Sensor Networks. International Journal of Computer Science and Network Security (IJCSNS), 2017, vol. 17, no 2, p. 194.
[63] SHAHZAD, Furrakh, PASHA, Maruf, et AHMAD, Arslan. A Survey of Active Attacks on Wireless Sensor Networks and their Countermeasures. arXiv preprint arXiv:1702.07136, 2017.
[64] SINGH, Rupinder, SINGH, Jatinder, et SINGH, Ravinder. WRHT: A Hybrid Technique for Detection of Wormhole Attack in Wireless Sensor Networks. Mobile Information Systems, 2016, vol. 2016.
[65] MULLA, Ms Raisa I. et PATIL, Rahul. Review of Attacks on Wireless Sensor Network and their Classification and Security. Imperial Journal of Interdisciplinary Research, 2016, vol. 2, no 7.
[66] DINKER, Aarti Gautam et SHARMA, Vidushi. Attacks and challenges in wireless sensor networks. In : Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on. IEEE, 2016. p. 3069-3074.
[67] YOUNIS ABDULLAH, Maan, HUA, Gui Wei, et ALSHARABI, Naif. Wireless sensor networks misdirection attacker challenges and solutions. In : Information and Automation, 2008. ICIA 2008. International Conference on. IEEE, 2008. p. 369-373.
[68] SAINI, Meenu, KUMAR, Rajan, et al. To Propose a Novel Technique for Detection and Isolation of Misdirection Attack in Wireless Sensor Network. Indian Journal of Science and Technology, 2016, vol. 9, no 28.
[69] ABDALZAHER, Mohamed S., SEDDIK, Karim, ELSABROUTY, Maha, et al. Game Theory Meets Wireless Sensor Networks Security Requirements and Threats Mitigation: A Survey. Sensors, 2016, vol. 16, no 7, p. 1003.
[70] KURBAH, Rangstone Paul et SHARMA, Bobby. Survey On Issues In Wireless Sensor Networks: Attacks and Countermeasures. International Journal of Computer Science and Information Security, 2016, vol. 14, no 4, p. 262.
[71] AGRAWAL, Dharma Prakash. Authentication, Encryption, and Secured Communication. In : Embedded Sensor Systems. Springer Singapore, 2017. p. 393-413.
[72] SARANYA, K., SATHYA, D., et KUMAR, P. Ganesh. A Survey On Intrusion Detection System In Wireless Sensor Networks. International Journal of Advance Research, Ideas and Innovations in Technology. 2017, vol. 3, no 1, p. 692-703
[73] MANOHAR, Ram Pradheep et BABURAJ, E. Detection of Stealthy Denial of Service (S-DoS) Attacks in Wireless Sensor Networks. International Journal of Computer Science and Information Security, 2016, vol. 14, no 3, p. 343.
[74] ACHARJYA, D. P. et AHMED, N. Syed Siraj. Recognizing Attacks in Wireless Sensor Network in View of Internet of Things. In : Internet of Things: Novel Advances and Envisioned Applications. Springer International Publishing, 2017. p. 149-172.
[75] MANSOURI, Djamel, MOKDAD, Lynda, BEN-OTHMAN, Jalel, et al. Dynamic and adaptive detection method for flooding in wireless sensor networks. International Journal of Communication Systems, 2017.
[76] SHAHID, Jahanzeb, SALEEM, Shahzad, et QURESHI, Muhammad Nauman. DOS Attacks on WSN and Their Classifications With Countermeasures-A Survey. NUST Journal of Engineering Sciences, 2016, vol. 9, no 2.
[77] CAPOSSELE, Angelo T., CERVO, Valerio, PETRIOLI, Chiara, et al. Counteracting Denial-of-Sleep Attacks in Wake-up-radio-based Sensing Systems. In : Sensing, Communication, and Networking (SECON), 2016 13th Annual IEEE International Conference on. IEEE, 2016. p. 1-9.
[78] GOPE, Prosanta, LEE, Jemin, et QUEK, Tony QS. Resilience of DoS Attacks in Designing Anonymous User Authentication Protocol for Wireless Sensor Networks. IEEE Sensors Journal, 2016, vol. 17, no 2, p. 498-503.
[79] GARCIA-FONT, Victor, GARRIGUES, Carles, et RIFÀ-POUS, Helena. Attack Classification Schema for Smart City WSNs. Sensors, 2017, vol. 17, no 4, p. 771.
[80] DINKER, Aarti Gautam et SHARMA, Vidushi. Attacks and challenges in wireless sensor networks. In : Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on. IEEE, 2016. p. 3069-3074.
[81] DENG, Jing, HAN, Richard, et MISHRA, Shivakant. Defending against path-based DoS attacks in wireless sensor networks. In : Proceedings of the 3rd ACM workshop on Security of adhoc and sensor networks. ACM, 2005. p. 89-96.
[82] DHAKNE, A. R. et CHATUR, P. N. Detailed Survey on Attacks in Wireless Sensor Network. In : Proceedings of the International Conference on Data Engineering and Communication Technology. Springer Singapore, 2017. p. 319-331.
[83] SARANYA, K., SATHYA, D., et KUMAR, P. Ganesh. A Survey On Intrusion Detection System In Wireless Sensor Networks. In the proceeding of the 2nd International Conference on Engineering Technology, Science and Management Innovation (ICETSMI- 2017)- India, January 2017. P, 158-165
[84] SAINI, Meenu, KUMAR, Rajan, et al. To Propose a Novel Technique for Detection and Isolation of Misdirection Attack in Wireless Sensor Network. Indian Journal of Science and Technology, 2016, vol. 9, no 28.
[85] ZHENSHAN, Bao, BO, Xue, et WENBO, Zhang. HT-LEACH: An improved energy efficient algorithm based on LEACH. In : Mechatronic Sciences, Electric Engineering and Computer (MEC), Proceedings 2013 International Conference on. IEEE, 2013. p. 715-718.
[86] RADHAPPA, Harish, PAN, Lei, XI ZHENG, James, et al.Practical overview of security issues in wireless sensor network applications. International Journal of Computers and Applications, 2017, p. 1-12.
[87] HE, Daojing, CHAN, Sammy, et GUIZANI, Mohsen. Cyber Security Analysis and Protection of Wireless Sensor Networks for Smart Grid Monitoring. IEEE Wireless Communications, 2017, vol. 24, no 6, p. 98-103.
[88] SINGH, Pooja et CHAUHAN, R. K. A Survey on Comparisons of Cryptographic Algorithms Using Certain Parameters in WSN. International Journal of Electrical and Computer Engineering, 2017, vol. 7, no 4, p. 2232.
[89] ZHANG, Ping, WANG, Shaokai, GUO, Kehua, et al. A secure data collection scheme based on compressive sensing in wireless sensor networks. Adhoc Networks, 2018, vol. 70, p. 73-84.
[90] CHELLI, Kahina. Security issues in wireless sensor networks: attacks and countermeasures. In : Proceedings of the World Congress on Engineering. 2015. p. 1-3.
[91] LI, Juan. A Symmetric Cryptography Algorithm in Wireless Sensor Network Security. International Journal of Online Engineering (iJOE)
[92] CHELLI, Kahina. Security issues in wireless sensor networks: attacks and countermeasures. In : Proceedings of the World Congress on Engineering. 2015. p. 1-3.
[93] KARDI, Amine, ZAGROUBA, Rachid, et ALQAHTANI, Mohammed. Performance Evaluation of RSA and Elliptic Curve Cryptography in Wireless Sensor Networks. In : 2018 21st Saudi Computer Society National Computer Conference (NCC). IEEE, 2018. p. 1-6.
[94] LI, Xiong, NIU, Jianwei, BHUIYAN, Md ZakirulAlam, et al. A robust ECC-based provable secure authentication protocol with privacy preserving for industrial internet of things. IEEE Transactions on Industrial Informatics, 2018, vol. 14, no 8, p. 3599-3609.
[95] IQBAL, Ummer et SHAFI, Saima. A Provable and Secure Key Exchange Protocol Based on the Elliptical Curve Diffe–Hellman for WSN. In : Advances in Big Data and Cloud Computing. Springer, Singapore, 2019. p. 363-372.
[96] BOJANOVA, Irena, BLACK, Paul E., et YESHA, Yaacov. Cryptography classes in bugs framework (BF): Encryption bugs (ENC), verification bugs (VRF), and key management bugs (KMN). In : Software Technology Conference (STC), 2017 IEEE 28th Annual. IEEE, 2017. p. 1-8.

[97] LIU, Donggang, NING, Peng, et DU, Wenliang. Group-based key predistribution for wireless sensor networks. ACM Transactions on Sensor Networks (TOSN), 2008, vol. 4, no 2, p. 11.
[98] JONES, K., WADAA, Ashraf, OLARIU, Stephan, et al. Towards a new paradigm for securing wireless sensor networks. In : Proceedings of the 2003 workshop on New security paradigms. ACM, 2003. p. 115-121.
[99] http://www.zigbee.org/.
[100] MUNIVEL, E. et AJIT, G. M. Efficient public key infrastructure implementation in wireless sensor networks. In : Wireless Communication and Sensor Computing, 2010. ICWCSC 2010. International Conference on. IEEE, 2010. p. 1-6.
[101] MALEH, Yassine et EZZATI, Abdellah. An advanced Study on Cryptography Mechanisms for Wireless Sensor Networks. arXiv preprint arXiv:1609.05323, 2016.
[102] OLIVEIRA, Leonardo B., ARANHA, Diego F., GOUVÊA, Conrado PL, et al. TinyPBC: Pairings for authenticated identity-based non-interactive key distribution in sensor networks. Computer Communications, 2011, vol. 34, no 3, p. 485-493.
[103] DIERKS, Tim. The transport layer security (TLS) protocol version 1.2. IETF RFC 5246, August 2008.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 244-254 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Current Trends and Challenges in Link Prediction Methods in Dynamic Social Networks: A Literature Review

Elfadil Abdalla Mohamed[*,1], Nazar Zaki[2], Mohammad Marjan[2]

[1]*College of Engineering and Information Technology, Ajman University, Ajman, UAE*

[2]*Dep. of Comp. Science and SW Eng., College of Info. Technology, UAEU, Al Ain, UAE*

ARTICLE INFO

Article history:
Received: 29 September, 2019
Accepted: 15 November, 2019
Online: 05 December, 2019

Keywords:
Link inference
Link prediction
Dynamic social networks

ABSTRACT

In more recent times, researchers have turned their attention to link prediction and the role link inference can play in better understanding the evolutionary nature of social networking sites. The objective of this paper is to present an in-depth review, analysis, and discussion of the cutting-edge link prediction methods that can be applied to better understand the development of social networks. The findings of the literature review reveal that there has been a steady increase in the number of published articles that present novel link prediction models that are designed to enhance the efficiency and accuracy of link prediction. In this paper, this most recent techniques to be proposed in this regard are compared and categorized, and features and evaluation metrics are presented for each approach. The results of the evaluation reveal that there are no complete or definitive methods available that can accurately and reliably be applied within different dynamic social networks to predict missing, emerging, and broken links within the network. The paper concludes by presenting potential future directions and recommendations for further studies

1. Introduction

This paper is an extension of work originally presented at the International IEEE Congress on Information Science and Technology [1].

Starting from the beginning of this century, social networks have significantly evolved, and the progression of social media platforms (Facebook, Twitter, LinkedIn, etc.) has been well documented in a significant number of publications. In the contemporary era of connectivity, the majority of organizations complement their traditional marketing strategies with digital campaigns that rely heavily on social media channels. Due to its increase in popularity, social media has become integral to organizations' advertising and marketing campaigns, representing a cost-effective and efficient channel through which companies target and communicate with members of a given audience.

Since their inception in the late 90s, the value of social networks (SNs) has dramatically increased to reach billions of dollars. As such, they represent an attractive investment proposition, especially in the marketing sector. Their rise in popularity and social and economic implications has also attracted significant research attention. One area of interest particularly notable in recent times is the prediction of missing links [2, 3].

1.1. Predicting Links in SN

By using graph theory concepts, SN can be represented as a graph of vertices and edges: G(V, E). The vertex V represents a user while the edge E represents a link. Figure 1 shows a simple representation of a social network.

The issue of predicting links in SN is concerned with the forecasting of possible connections and interactions that can be observed between various members within a certain network. As explained in [4], considering an SN, G, at a certain time, t, the objective of link prediction here is to predict possible new links or the breaking of existing ones at a later time, t~. Several researchers have expressed an interest in predicting existing, expected, and missing links by developing a range of different methods. Despite the fact that several methods have been proposed, however, [5] indicate that none of them are reliable in effectively predicting the missing links.

An important feature of link prediction in SN is concerned with identifying missing links. Having the ability to predict missing links would provide us with clues of how SN evolves within a

* Elfadil Abdalla Mohamed, Email : elfadil.abdalla@ajman.ac.ae

range of settings. Predicting missing links has important connotations, for example, in the academic world, where doing so would help to identify possible academic collaboration in certain fields of interest [6, 7].

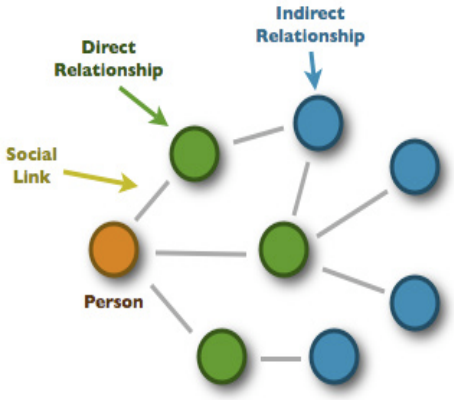


Figure 1: Simplified representation of SN

A further area where social link prediction could be of use is within criminal investigation; for example, similarity-based link-prediction methods are viable for identifying links between members of criminal networks. Identifying possible missing links in criminal networks can be achieved by exploiting node similarity in noisy or incomplete situations [8]. Predicting broken links is also of paramount importance within commercial settings, as marketing, customer service, and customer experience strategies based on link prediction can maintain customer loyalty. In bioinformatics, biology, and healthcare, the prediction of missing links can assist organizations in locating the relevant specialists who are able to receive future referrals, while within gene expression networks [9] they can be employed to better understand protein-protein interactions. Elsewhere, as explained in [2, 10], within security-related networks, link prediction can be used to detect suspected communications that countries are more likely to find harmful.

Social networks contain significant amounts of data; as such, it is not possible to collect all the information connected to a user's relationships. Link prediction, therefore, represents a viable means by which it is possible to apply an understanding of known relationships between users to estimate unknown relationships.

The existing literature highlights a continual increase that has occurred in the number of link-prediction methods that are capable of predicting SN links. This vast rise in the number of such methods is exponential to a large number of publications in renowned journals or conference proceedings that either propose a new method of predicting links or else to improve the accuracy and efficiency of an existing approach. Figure 2 depicts the number of articles published in Scopus indexed journals or conference proceedings. As the graph reveals, there has been a steady rise in the number of articles published that have focused on link prediction practices. Subsequently, many solutions have been introduced to handle the problem of accurately and efficiently predicting the missing links.

A link-prediction method that exploits the information diffusion feature was discussed in [11]. This study demonstrated the correlation between information diffusion processes and the creation of new links. The proposed method managed to enhance the accuracy of link prediction by virtue of information diffusion and was tested on a Sina Weibo dataset. The experimental data revealed that the proposed method outperformed approaches that rely on topological features alone.

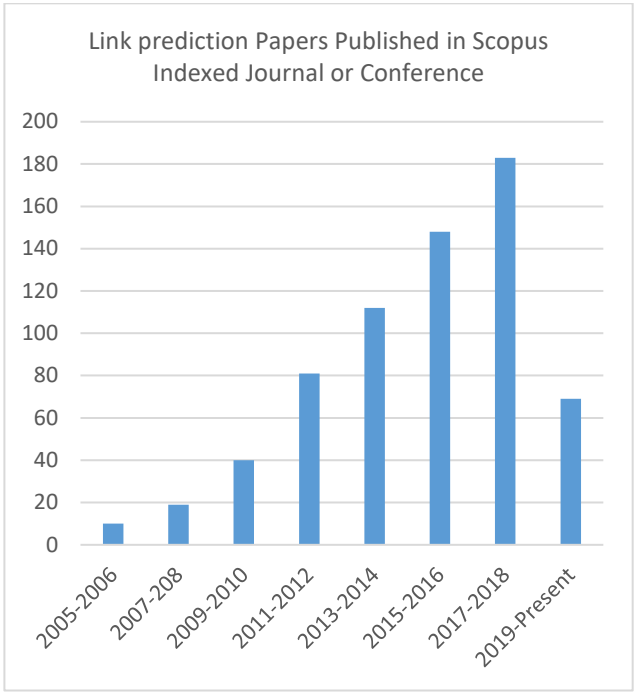


Figure 2: Papers related to link prediction published in Scopus-indexed journals or conference proceedings

Several link prediction approaches that are based on a deep belief network were presented in [12]. The results obtained from the experiments on datasets collected from different sources revealed that the methods effectively predicted link values and exhibited a remarkable generalization capability among the studied datasets.

Due to the vast amount of data in social networks, achieving efficiency within link-prediction methods represents a significant challenge. To address this problem, [13] proposed two algorithms for link prediction. The proposed algorithms solved the efficiency problem by adopting low-rank factorization models, while also proving very efficient compared to other methods. As such, the study represents a significant step forward in the challenge of developing an efficient link-prediction model.

Another important problem that needs to be addressed concerns the accuracy of link prediction approaches. [14] developed an innovative link-prediction method that aimed to improve the accuracy of link predictions. The experimental findings revealed that the proposed approach outperformed many other methods in terms of accuracy and scalability and the associated runtime was significantly less than that observed in previous studies. However, although the proposed approach did enhance link prediction accuracy, it remains untested on large network data, which may see its efficiency become undermined.

Improving the accuracy of link prediction can also be achieved using methods that predict links in uncertain frameworks. [15] proposed a new approach for addressing the problems of link

prediction in the context of an uncertain framework based on the theory of belief functions. Results obtained from the experimental work of [15] outperformed traditional approaches.

Another problem related to link prediction in signed SNs concerns the sparsity of data. To alleviate this problem, [16] proposed an innovative approach that is capable of exploring the personality of the user using social media. The experimental results obtained indicated that a complementary relationship exists between the signed link prediction problem and personality information.

An interesting state-of-the-art method for facilitating link prediction is the use of the temporal regularity in interpersonal communication as a means of prioritizing weighted edges in network graphs [17]. Compared with other methods, this method predicts links even if there is a scarcity in the number of edges needed for analysis.

A contrasting similarity-based link-prediction method, based on fuzzy link importance, was presented in [18]. The method performed well, using two strategies to achieve its objectives. Firstly, for the selection of the neighbor, the distance between nodes was used. Secondly, to find the relevant link, the fuzzy link importance was employed. By using these strategies, the method has obtained sound results.

An approach aiming to solve the problem of link prediction problems in large complex networks was proposed by [19] following two steps. In the first step, the complex network is transformed into a different simple network, before the second step uses probability to predict the possible links. The experimental test results indicate that the method achieved promising results.

Methods for predicting temporal links in temporal networks have been introduced. For example, [20] proposed an innovative approach based on a semi-supervised learning network with the aim of predicting current and future links. The approach was tested using real data. The results obtained from the experimental test demonstrated that the approach was very effective and highly scalable.

A myriad of similarity-based methods has been introduced to address link prediction problems. However, methods of this nature require substantial data to work effectively. Recent methods have been introduced to alleviate such problems, and approaches based on similarity indices have accomplished considerable efficiency. For example, [21] proposed a method that uses a clustering coefficient index found to outperform many existing approaches. The improvements the approach exhibited were attributed to the use of an index of clustering.

Within the context of multiple networks, the link prediction problem was discussed in [22]. In terms of accuracy, the supervised learning method achieved 92.5%, which is very promising. The unsupervised method achieved a resounding accuracy result of 97% using normalized discounted cumulative gain.

Another interesting method proposed as a means of addressing the problems of similarity-based link prediction approaches was put forward by [23], which successfully managed to address the issue of lack of efficiency. The experimental results revealed that the LDR index enhanced the prediction performance in undirected networks. However, the researchers highlighted a need to investigate the method's application further within directed and signed networks.

Many researchers have achieved remarkable improvements in link prediction within the context of social networks. However, studies that address trusted link prediction are very few. Recently, [24] proposed a method that sought to forecast links by using the user's most important features. The experimental results revealed that the method outperformed many existing approaches in terms of effectiveness. Moreover, the quality of community detection was improved.

In the past, research that investigates issues related to link prediction in signed networks has been very scarce. Only recently have we witnessed an increase in the number of studies that discuss issues related to link prediction in signed networks. For example, [25] proposed a fresh approach with experimental results that show a high quality in terms of the accuracy and efficiency of link predictions.

Many further new methods have been introduced with the aim of improving link prediction accuracy, for example, [26] has proposed a method that achieves good accuracy using associated degrees.

A more interesting method for link prediction was proposed by [27], which employed the level-2 node clustering coefficient. As previously described, most link-prediction methods suffer from either lack of efficiency or accuracy. Plus, increased efficiency can lead to a reduction in accuracy. [27] suggested a method that was tested experimentally over 11 real-world datasets. The obtained results indicated that the method outperformed many of the state-of-the-art existing methods in terms of accuracy.

In dynamic SNs, categories of users exert social influence on other members within the social network. These individuals are commonly referred to as influencers. Influential users play a fundamental role in digital marketing [28]. As such, predicting links to and from influential users is of paramount importance for many organizations. [29] introduced a method aiming to predict influential links. The approach was tested experimentally and the test's findings revealed that the proposed approach outperformed comparable link prediction metrics. Moreover, the link prediction performance was improved in comparison to that associated with classical topological metrics.

Low link prediction accuracy remains a major challenge that requires attention. A myriad of methods and algorithms has been introduced by many research field scholars. A current approach that aims to address the problem of low link prediction accuracy, which is based on matrix factorization, is proposed by [30]. The experimental results of the method indicate that it achieves a higher prediction accuracy than existing methods.

Despite the fact that matrix-factorization-based methods have achieved good security, such methods suffer from the problem of having to build the adjacent-matrix. [31] has developed a technique to address the problem of matrix-factorization-based methods by fusing the adjacent matrix and some key topological metrics in a unified probability matrix. The [31] models were experimentally tested with real data, with the results of the

experiment revealing that the proposed models achieved an impressive link prediction performance.

Another important and interesting method that has achieved great link prediction accuracy was put forward by [32], which used H-index and the influence of degree. The method achieved remarkable link prediction accuracy by utilizing information about the endpoint node.

Among the methods that have been developed to utilize real node influence and improve link prediction accuracy is that of node ranking [33]. This method uses the concept of node ranking to improve link prediction accuracy.

Even though there are some problems inherent in similarity-based methods, they have achieved considerable results in terms of accuracy and efficiency. However, further such improvements will be required. A method that is based on the kernel graph, which utilizes the structural information extracted from a signed social network, was proposed in [34]. This method involves initially generating a set of subgraphs with different strengths of social relations for each user. Having been tested with real data, the experimental test results revealed that the method achieved considerable link prediction performance on the two types of positive and negative links. Moreover, the accuracy and F1-Score exceeded that of existing methods, indicating that there is room to improve accuracy.

More methods that use a similarity index for link prediction have recently been introduced. A method that took the attribute similarity between the node pair into consideration, named attribute proximity, was discussed in [13]. The experimental test results showed that it achieved higher accuracy than approaches that didn't take node attribution into account.

1.2. Link Prediction in Dynamic SNs

Dynamic networks can be used to describe members that exhibit varying dynamics over time [1]. In such a network, a new member joins, existing members may leave the network, and members create or break relationships [46]. As a consequence of this phenomenon, the network can expand or shrink. For this reason, there is a compelling need to take dynamic changes into consideration when developing approaches that can accurately predict new, current, or missing links. Link prediction in temporal SNs is more challenging due to the continuous network changes. The dynamic changing nature of temporal SNs will result in the emergence of different types of sub-graph. Many applications have attempted to exploit this nature of networks, for example, countries seeking to beef up their security can use such features to predict criminal links, and online recommendations [4]. Biological networks can also utilize temporal networks, with a vast number of applications having been proposed that do so, [57] developing innovative methods for detecting protein complexes. The method proposed by [57] exploited the dynamic nature of protein, while [35] has developed a model that detects the progress of a community over time by utilizing the temporal nature of SNs. [36] discussed an interesting method that predicted links in a dynamic SN based on three metrics. The proposed method was experimentally tested on DBLP data and the results obtained indicate that the method achieved superior results to alternative approaches. The predicting of future links in social networks using the proximity of node is discussed in [37]. In this approach, future interaction can be obtained from network topology alone.

The structure of social networks is rapidly evolving. To leverage this evolving structure, [38] proposed a method that exploited characteristics calculated by a known time prediction of measures computed using a pair of nodes. The method was tested using real data. The experimental results obtained revealed that the approach achieved high accuracy.

The influx of methods that are capable of predicting future relationships in dynamic SNs continues to increase. For example, [39] introduced a method that utilizes learning automata for link prediction. In the same direction another new method proposed by [40] uses learning automata to predict links in stochastic SNs. The method was tested using data collected from social networks and has achieved sound link prediction accuracy. One of the interesting features of this is that it deals with online stochastic SNs that can encounter complex online variations.

Clustering, one of the prominent data-mining techniques, has also been employed for predicting future links in dynamic SNs. An interesting method that uses the clustering approach for predicting future links is proposed by [41], while another method is discussed in [42]. These two methods have obtained sound results in comparison with existing methods.

Deep learning is one of the data-mining techniques that has received more interest from researchers working in the field of data mining and machine learning. The nature of this technique makes it suitable for use in dynamic SNs. [1] discussed a number of techniques that predict links using deep learning. Also, a viable approach that successfully managed to predict links using machine learning is discussed in [43].

Some link-prediction methods suffer from a lack of precision. However, some researchers have attempted to alleviate such a problem by proposing different types of techniques. For example, [44] proposed an approach that utilizes community information to improve prediction accuracy. Taking the same approach, [45] introduced a technique that exploited the structure of community information to predict new edges and hence improve link prediction precision.

One of the most important features of online SNs is their rapid change over time [46]. Currently, most of the methods used for predicting links depend on structural SNs. There is a compelling need for methods that are capable of predicting links in dynamic SNs. Moreover, the required methods should also possess high link prediction accuracy and efficiency. Organizations intend to exploit online SNs for a number of applications and require information on state-of-the-art methods that can be used to predict possible links. The main objective of this paper is to discuss and analyze the most prominent link-prediction methods that are able to accurately and efficiently predict links in dynamic SNs.

This paper is organized as follows. Section 2 discusses the problems of link prediction. The next section discusses the strategies used by link-prediction methods to solve the problem of link prediction. This is followed by discussing and analyzing existing work. The paper wraps up with conclusions and future work.

2. Problems of Link Prediction

A simple question and answer summarizes the problem of link prediction. As discussed in [47], if we have SN at a certain time, t1, after a period of time, t2, can we predict a possible link between members of the SN? Many methods have been introduced that use structural SNs but, in real life, SNs are dynamic. In this section, we try to provide an overview of the problems of link prediction in dynamic SNs.

2.1. Link Prediction in a Temporal Network

A temporal network is a network that evolves over time. For predicting the emergence of links in such a type of a network, time factors are crucial [48]. Details on how links can be predicted using a temporal network can be found in [49]. The main problems and challenges of predicting links in a temporal network that require addressing are related to the changing nature of the network over time. Many methods and techniques have been proposed to address this problem. For example, [48] introduced a technique to exploit the general network structure and managed to successfully predict links in the process. [5] proposed a technique using graph theory to predict links, and [50] used community structure information to do so. An efficient method proposed by [51] produced far better results than many contemporary temporal link-prediction methods.

In [52], the authors explored methods of inferring links in a special type of network; ego-networks – where there is a scarcity of information about neighbors. The proposed methods introduced numerous approaches to retrieve information about communities. In doing so, they have improved the prediction of links compared with other methods that use structural SNs. An interesting method that determines the edge weight based on a calculation of the scores using spectral analysis is given in [17]. The method achieved good prediction results.

An effective and scalable method for predicting links within a real-world temporal social network was presented in [20]. [53] proposed a method that takes into consideration the time features of the network. The method captured the importance of timestamps in the interaction between network nodes and the experimental results showed significant improvement of link prediction. In the context of a temporal network, [54] proposed a method that predicts links by using a cross-temporal concept, which involves inferring the nodes at different time intervals. The method was tested experimentally using real-world data and the results showed sound improvements in the link prediction accuracy.

2.2. Link Prediction in Heterogeneous Networks

In SNs, there are two major types of networks: homogeneous networks and heterogeneous networks. The homogeneous networks assume that all the nodes and edges are of the same type while, in heterogeneous networks, there is variation in the nodes and the links between them. The majority of the existing link-prediction methods deal with homogeneous SNs. In the real world, the majority of SNs are heterogeneous with different types of nodes and edges that pose many challenges requiring intervention. For addressing such problems, many new link-prediction methods have been introduced. For example, the work of [55, 56, 5, 57] and recently [58].

2.3. Link Prediction with Active/Inactive Links

Interactivity in SNs is generally represented by sending e-mails, receiving phone calls, commenting on messages, etc. Historical information related to interactivity between nodes in the networks, such as the timestamp, is likely to improve the accuracy of link prediction. Using the temporal features of dynamic SNs has been proven to improve accuracy too [4].

2.4. Link Prediction Scalability

The era of big data has emphasized the importance of efficiency for link prediction algorithms. The link prediction algorithm scalability remains questionable unless the algorithm is tested and evaluated with a huge amount of data. However, most of the current existing link-prediction methods have been tested and evaluated using limited datasets, which make it harder to ensure their scalability. The issue of algorithm scalability remains one of the challenges that need to be addressed. Attempts were made by [59] to develop a scalable link prediction algorithm. The developed algorithm managed to predict links using features of endpoints and neighborhood, adopting the locality-sensitive hashing algorithm to enhance the scalability so that the proposed approach could effectively predict links in large networks spanning long-term sequences.

3. Link Prediction Strategies

This section discusses the strategies used for link prediction, which include similarity-based methods, maximum likelihood methods, clustering-based methods, probabilistic models, fuzzy link methods, and matrix factorization-based link-prediction methods.

3.1. Similarity-based Strategies

Similarity-based link-prediction methods are among the first and simplest methods used in link prediction. The idea behind this class of methods is that each node pair – say n1 and n2 – is given a score that reflects the similarity between n1 and n2. The algorithm then ranks the pair of nodes based on their score, and the node pair with the highest score is most likely to have a link between them [60]. Nodes show more similarities if they have shared neighbors.

There are some problems and challenges associated with the link-prediction methods that use a similarity score. As discussed in [18], predicting the link is based on computing the rating and the selection of the neighbors by using similarities between the pairs. If the algorithm fails to find enough information on the ratings, there will be problems with computing the similarity. Moreover, the accuracy of link prediction will be negatively impacted by the number of neighbors.

Similarity-based link-prediction methods are not confined to structural SNs. In dynamic SNs, a number of methods have also been introduced. For example, [4] and [50] proposed an interesting method that employs similarity indices in dynamic SNs. The model proposed in [4] is based on the Covariance Matrix Adaptation Evolution Strategy. Plus, a modified similarity-based link-prediction method was discussed in [61]. The approach was experimentally tested and evaluated using real-world data. The results obtained indicate that the method outperformed existing similarity-based link-prediction methods in terms of accuracy. Two improved algorithms, based on the similarity method that applies the network topology sufficiently, were presented in [62]. The experimental results of the two improved algorithms reveal that the prediction performance was far better than the existing traditional one.

Some efficiency improvements have been made to the similarity-based strategy. For example, [23] proposed an algorithm that improved computation efficiency. The improvements were made by the algorithm with the virtue of using linear computation. Further improvements in terms of effectiveness, efficiency, reliability, and strength were achieved using the method introduced by [24]. This approach made use of several user features to predict trusted links. Additional improvements in terms of the accuracy of link prediction were achieved by [26]. Their method used strategies that depend on the depth of the path passing from the source to destination nodes and the associated degree.

Accuracy, which is consistently used to evaluate link-prediction methods, represents an important consideration when developing and implementing any method that can be used for predicting links. A modified similarity-based method employing graph theory was proposed by [34]. Their method achieved significant improvements in the accuracy of predicting links compared with existing methods, thereby demonstrating that similarity-based methods continue to play a significant role in link prediction. The accuracy of similarity-based link-prediction methods was further improved by [63]. Their method achieved a remarkably high prediction accuracy by using a similarity index called attribute proximity. The authors concluded that the higher the similarity between the topological neighborhoods of two nodes, the more likely that a link will emerge.

3.2. Maximum Likelihood Algorithms

Maximum likelihood algorithms for link prediction is a concept applied to determine the parameter distribution of the network structure. In link prediction, the assumption made here is to organize the network structure and then maximize the likelihood of the observed structure. Based on this concept, a number of methods have been developed. For example, [55] employed the concept of maximum likelihood to predict links in SNs. With limited datasets, the method obtained good results.

Link-prediction methods based on maximum likelihood face practical applications related to the time taken by the algorithm to converge, especially if the dataset is very large, implying that the algorithms are not scalable.

3.3. Probabilistic Models

An interesting class of methods used to predict SN links is based on a probabilistic concept. This approach can be applied to both static and dynamic SNs. In dynamic SNs, [64] employed a probabilistic concept to deal with the complexity of a non-linear dynamic created by the data features. The model managed to predict links with good accuracy. Additional methods were proposed using this concept to predict links such as [55, 56 , 5]. Some methods combined the features of graph theory and probabilistic concept to predict links in SNs [65]. This hybrid method is very scalable due to the innovative approach used to approximate the probability of links.

In [66], the authors proposed an algorithm framework that combines probability with the Hamiltonian structure to predict links in SNs. The algorithm was tested and evaluated experimentally using numerical simulation data. The results obtained from testing the algorithm showed that it achieved a very high accuracy compared with the best available link-prediction methods. The algorithm also successfully managed to identify and uncover the missing links.

Despite the fact that link-prediction methods based on probabilistic features have achieved sound results in predicting links in SNs, the main problems of these methods center around efficiency. The computational time needed by the method is very high, so there is a need for reducing this reliance.

3.4. Clustering-Based Link-Prediction Methods

Clustering-based link prediction approaches have played a considerable role in enhancing the efficiency and accuracy of links inferencing. These methods have been discussed in several studies, for example, [11, 67, 21, 27].

In [11], the method achieved better improvements in the accuracy of links forecasting as the number of clusters grew, whereas [67] proposed a cluster-based method that used clustering information to predict the missing links. Their method was tested on three large-scale networks, with the experimental results revealing that it outperformed other approaches and, more interestingly, that information about link clustering has improved the accuracy of link prediction. [21] used an index that took into consideration more information related to the structures provided for link forecasting to enhance its accuracy in comparison to alternative indices. The accuracy of the method proposed by [21] was compared with twelve representative link-prediction methods, and the findings revealed that it exceeded their performance.

The approach proposed by [27] first involved extracting similar neighbors and grouping them into clusters, then computing the coefficients of the cluster up to two common levels. Their method exceeded the performance of all the alternative baseline methods. The only method that outperformed theirs is Node2vec, albeit with medium accuracy.

3.5. Fuzzy Link-based Methods

Link prediction based on the fuzzy concept has only recently been applied. For example, [18] proposed a model based on this concept with the proposed model, when tested and evaluated on real data, achieving considerable accuracy improvements. [68] proposed two methods that employ the concept of fuzzy-link prediction. The first approach used the concept of a fuzzy soft set, while the second employed the Markov model concept to improve the efficiency of link-prediction methods. They claimed that their models achieved better prediction than the existing approaches.

3.6. Matrix Factorization Methods

The simple idea behind matrix factorization is to decompose a complex matrix into a simpler one to make it possible to compute more complex operations. Numerous link-prediction methods based on matrix factorizations were proposed that have effectively improved performance in terms of the accuracy and efficiency of link prediction. For example, [31] propose a framework that can be used for link prediction incorporating the matrix factorization concept. [69] discussed an interesting link-prediction method that uses the concept of matrix factorization. The proposed method produced better results in terms of link prediction accuracy when compared with other methods. A recent study based on matrix factorization was discussed in [30]. Although the method of [30] achieved high prediction accuracy, the data was relatively limited as it was confined to sales datasets. As such, further evaluation involving more complex datasets is required to confirm its level of prediction accuracy.

4. Analysis and Discussion of Existing Methods

In the previous sections, we have discussed the challenges and problems of predicting the emergence of new links using a special type of SN; namely, dynamic SNs. Furthermore, we have also discussed the different types of link-prediction methods proposed by many scholars working in the SN field. This section discusses the most current solutions introduced by researchers in the field of predicting links in SNs. The main topics to be discussed in this section start with the summary of the proposed solutions for addressing links prediction problems, followed by datasets used, the main features of link prediction, the techniques used in link prediction, and concluding with the evaluation and accuracy measures. A summary of the solution's evaluation results is depicted in Table 2.

4.1. Summary

Once online SNs had become an important platform for the exchange of a huge amount of data between users, their existence attracted a significant number of scholars intending to study how these networks evolve over time. The emergence and deletion of links help researchers to understand the dynamic nature of SNs [2].

In the past two decades, many link-prediction methods have been introduced, implying that the problem is not new. However, the new and most challenging SN-related task concerns the forecasting of links in dynamic SNs, which are characterized by growing or shrinking networks. The majority of the work that has been performed in the area of dynamic SNs to date has been conducted with the acknowledgment of the importance of a dynamic nature within the network. Many scholars have taken into consideration the changing time features of the SN by using past information related to the node transactions, while others used the temporal feature of community structure [50]. These approaches forecast the future importance of a node based on eigenvector centrality. Subsequently, this view of future importance is used to predict links.

Some researchers have tried to develop new techniques for predicting future links using historical data. Their main objective is to improve the efficiency and accuracy of the methods applied. A model for link prediction in dynamic SNs using deep learning is discussed in [1]. Another model for link prediction based on a triad transition matrix, using statistical data, is given in [56].

The model discussed in [55] predicts future links by using the Euclidian latent distance. Nodes that exhibit closeness in the Euclidian distance are more likely to emerge as links. [4] proposed a model depending on the local information of nodes to predict future links in the dynamic SNs, while [5] introduced a model that uses graph theory to predict links by including temporal features.

In [17], it is introduced a method for link prediction based on graph theory. The method used the spectral analysis technique to calculate the score for each edge in the network graph. Based on this score, the edge weight will be determined. The method was tested and evaluated using real-world data, with the obtained results indicating that the method has enhanced the prediction of links. More interesting, however, is that the edges used in the analysis were very limited.

4.2 Datasets Issues

Most of the research conducted for the sake of developing link-prediction methods involves the use of some form of data that is either artificial or real world, as with [64] and [55]. The major problem concerning the datasets used for testing and evaluating the link-prediction methods is the absence of standardization [2]. Researchers have used numerous SNs of various sorts and sizes according to their proposed model and there are several dataset types available for researchers to use. The majority of networks used in the existing literature include collaboration networks, online SNs, generated datasets, and friends and family e-mails. [55, 56, 5] have used collaboration SN datasets. [64] employed two biological SNs. Most of the datasets used are small in size, with the exception here being [4] who employed a moderately large size of the online SN (Twitter) for the development of their method. For testing and evaluating their method, [50] used five email SNs. [56] employed a merging of two datasets; namely, scientific collaboration networks and email. Table 1 shows a summary of the datasets used by various researchers surveyed in this paper.

Table 1: Datasets used by researchers [1]

Datasets	No. of nodes	No. of links
Irvine Msg	896	2,201
Enron	200	714
News Words	503	15,638
Manufacturing Emails	167	3,250
Robot	352	1,611
Twitter (RRNs)	$5.71*10^6$	$49*10^6$
DBLP	437,515	1,359,471
astro-ph	55,233	644,496
Friends and Family Emails	1,130	1,344
Math	24,818	506,550
Ask	159,316	964,437
Super	194,085	1,443,339
Stack	2,601,977	63,497,050

From the literature that we surveyed, we found that the choice of suitable dataset depends on the judgment of the researchers who want to develop the link-prediction method and evaluate its performance using the collected data. Furthermore, the datasets used in many applications are synthetic, artificial or generated [64] [55] which might raise some concerns about the results obtained using such a type of dataset. Moreover, the size of the data used in testing the performance of the methods is not large. Another problem regarding the data is the nature of the datasets which, in many cases, is static and needs to be used to test the method based on the dynamic network. Some researchers have attempted to address this problem by including features of a temporal network [64, 55, 56, 5]. As most SNs grow and shrink over time, indicating that these networks are dynamic, it will prove more interesting for link-prediction methods to be applied in these dynamic SNs rather than structured ones. To judge the practicality of such methods, the datasets used should be real and applicable for link predictions. As many studies indicate that the datasets used for evaluating the developed methods are either static, artificial, or with a sample size, these issues represent great challenges that need to be addressed. These challenges point to the important need for setting database standards that will be used for evaluating the performance of all these methods.

4.3. Features Used in the Link Prediction Process

The two main features used to predict links in dynamic SNs are node attributes information and network topology. Both features

have been used by numerous studies to predict links in dynamic SNs. For example, [55] developed a model that employed the latent space concept and used the local topologies feature. The model predicts links based on past information, assuming that if there are two nodes, n1 and n2, which have links in the time, t-1, then mostly likely they will be gaining a link at a time, t.

The model proposed by [4] employed the two predominant features used to predict future links in dynamic SNs; namely, the network topology and node attributes. A test of this model using a vast amount of data indicates a high convergence with astonishing precision. Unlike the majority of the techniques proposed to detect dynamic community, [4] proposed a model named Hierarchical and Overlapping Community Tracker (HOCTracker) that is capable of detecting the progression of an overlapping community in dynamic SNs. The results obtained from the experimental test of the model indicate that the community structure detected is far better than the best available methods.

The model proposed by [56] uses the concept of a triad transition matrix to discover the dynamic pattern of networks. The main feature used is a local topology network and observations of the recorded network history. The model merges statistical characteristics with the topology of the network. This merger resulted in a method with the strength to perform better in detecting the evolution of a network. The model introduced by [5] incorporated the available past information about links on the current SN state. The model shows that including a timestamp of the historical interaction improves the accuracy of predicting new links. The technique used by the model to predict links is a joining of the graph with the temporal information included in the evolving SNs.

4.4. Techniques Used in Link Prediction

Link-prediction methods employ a variety of approaches and strategies in the form of a cluster-based, similarity-based, or fuzzy link, etc. The most appropriate prediction technique is selected on the basis of the features used in the link-prediction method. In similarity-based models, the method predicts the links among nodes based on a similarity score computed among the pair of nodes. A pair of vertices with high similarity value is most likely to foster a relationship or link in the future. Methods based on maximum likelihood predict future links between two nodes by identifying a centric node. If the likelihood that the two nodes fall in the radius of that centric node is high, then there is a possibility that a link might emerge between them. Methods based on a probabilistic approach use probability to predict the future links between two nodes in the dynamic SNs. The methods predict the links by calculating the probability of the edge weights of the two nodes and the probability of the neighboring nodes. The clustering-based link-prediction methods use a node clustering coefficient to predict the link. The technique first computes the clustering coefficient then uses it to predict the future link [27].

The research of [4] is focused on a node's local feature, for which they use the similarity index to predict the emerging link. The work of [55] used an approach that ranks the likelihood of two vertices expected to be linked based on the frequency of joining in the test set.

The model proposed by [56], which is based on the Triad Transition Matrix, relied on the probabilistic approach in that it contained triad transition probabilities in the network. The local probabilistic model was extended by [5] to predict links using the timestamp feature and demonstrating the incorporation of link weight into the prominent link prediction approaches. [50] model integrated various kinds of information into the SNs, such as node centrality and temporal information.

Another model that relied on the probabilistic approach uses Conditional Temporal Restricted Boltzmann Machine (CTRBM), described by [64], to forecast emerging links on individual transition variance and the influences of local neighbors.

4.5. Methods Evaluation

This section discusses the diverse evaluation measures and metrics employed by the different types of methods proposed by various researchers. The evaluation in terms of method accuracy and efficiency represents an important factor for its applicability and acceptability. To date, there are no common metrics and measures that can be employed to evaluate these methods. This lack of common evaluation metrics and measures, coupled with the absence of standardized datasets, entails that it is very hard to assess each solution and, as such, we cannot present a reliable and valid conclusion as to which is the best in terms of link forecasting in dynamic SNs.

The methods explored in this study can be evaluated either by two machine-learning measures; namely, receiver operating characteristics (ROC) [70] and the area under ROC (AUC) or other measures. The interpretation of the values of these measures depends on the method, as each method has features that are different from the others.

For example, when evaluating two models that have an AUC score close to each other, it will be very hard to determine which one is the best to select. As such, [64] employed a measure that computes the sum of the absolute difference to distinguish between various types of models. The methods discussed in [4, 55, 15, 17] employed ROC and AUC for evaluation.

Researchers in [56] employed another means for evaluating their methods in terms of efficiency. They have used two evaluation techniques, with the first one being known as average normalized rank (ANR), and the second known as discounted cumulative gain (DCG), which is employed for node ranking. ANR was employed to directly pinpoint the location of the important item in the ranking. In comparison, [20] used two measures, F1Score and Kendall's Tau Coefficient (KTC), to evaluate their methods.

From the surveyed and discussed methods, we found that each researcher has typically selected the evaluation metric and measure that best suits their model. As a result of this, while every researcher claims to have developed a model that offers superior performance to the alternative methods, this performance was measured in a testing environment directly tailored to the method being assessed.

This variation in the method evaluation indicates that it is very hard to tell which model is superior to the others.

5. Conclusion and Future Recommendations

Link prediction within the context of social networks is by no means a novel research topic. However, the greatest challenge associated with predicting new or missing links in dynamic SNs characterized by ongoing evolution is yet to be adequately addressed.

Table 2: Summary of Famous Link Prediction Methods

No.	LP Strategy	Dataset	Proposed Model (Method)	Features Used	Evaluation and Accuracy Measures	Remarks
[64]	Probabilistic	Synthetic data A and B (10,000 nodes each) Robot.Net (2,483,776 nodes) Two biological datasets (Control and Exposure 13,483,584 nodes each) from Stevenson	Deep learning framework (ctRBM) + can quickly learn	Node attributes and network topology	ROC and AUC curves Sum of Absolute Differences (SumD)	Reliable for handling uncleaned data
[4]	Similarity-based	51 million tweets Twitter reciprocal reply networks over 81 days	Covariance Matrix Adaptation Evolution Strategy (CMA-ES)	Node-specific features and network topological similarities	ROC and AUC	Does not require parametric thresholds. The linearity of the model is one of its weaknesses. Can handle different types of network
[50]	Similarity	Irvine Message Enron Email Infectious SocioPatterns News words Nodobo	Integrated Time Series Model (ITM) that combines topological (community and centrality) and temporal information	Network topology, community structure, and node centrality feature	AUC	More realistic for dynamic networks
[56]	Probabilistic	Two email social network datasets Enron dataset Wroclaw University of Technology	Triad Transition Matrix (TTM) comprising the probabilities of transitions between triads found in the network	Network Local topology observations of the recorded network past information	Compared to preferential attachment (PA) and common neighbors (CN) predictor	Data was pre-processed
[55]	Maximum likelihood and probabilistic	NIPS co-publication data + synthetic data generated by a model	DSNL model handle change of friendship	Local topologies (represent each entity with a location p in the latent space)	AUC & ROC	Little data with only around 11,000 nodes. Cannot handle large data
[5]	Probabilistic	DBLP and astro-ph of ArXiv.	Graph-based link prediction techniques based on the local probabilistic model	Time-based edge weights derived from temporal aspects	Discounted Cumulative Gain (DCG) and Average Normalized Rank (ANR)	Cannot handle many types of networks
[17]	Temporal Method	Enron E-mail, Irvine SNS, Friends and Family SMS	Temporal regularity in interpersonal communication	Node-pair records extraction and time-domain signal generation	Area Under the Curve (AUC)	There is a need to take into consideration the directions of interpersonal communication
[18]	Fuzzy Link	100,000 ratings and 1,300 tag applications	Distance between two different objects	Links among items, fuzzy link importance	MAE and RMSE	There are limitations on the procedure used for evaluating the method
[20]	Temporal Network	Data obtained from (SNAP) Math, Ask, Super, Stack	(Cox PHM), Game theory	Topological features	Average F1Score, Kendall's Tau Coefficient (KTC)	The first method achieved good accuracy

This paper examines the most current and innovative methods for predicting links in dynamic SNs. Each method explored and discussed in this paper exhibited some sort of good performance, such as accuracy, efficiency, or scalability, in respect of one SN type. A variety of models were explored in this paper. For each explored model, the paper discussed the prediction strategy, the datasets used, the type of features, and the evaluation measures used.

One vital feature of dynamicity concerns the temporal aspects of SNs, which is something that many researchers have considered. Some emerging methods have made great strides in terms of link prediction accuracy and efficiency, including the novel methods that have used different topology and harnessed both supervised and unsupervised techniques. Some methods, such as temporal-based methods, are highly scalable in terms of their ability to predict links [20].

This rigorous analysis of the emerging solutions indicates that there is no one complete method available that predicts emerging or lost links in dynamic social networks to a high degree of accuracy, efficiency, and scalability. Continuous efforts are needed to develop an innovative model that is capable of handling different types of SNs.

Despite the influx of research that aims to predict links in dynamic SNs, we still believe that there is a need for further research to address the problem of the absence of complete research that exhibits high accuracy, high efficiency, and is very scalable as well. Further improvements in accuracy, efficiency, and scalability are very important due to the huge data generated from dynamic SNs.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to acknowledge the assistance and support provided by Ajman university library and United Arab Emirates University library for procuring most of the papers and the books used in this article.

References

[1] Mohammad Marjan, Nazar Zaki, and Elfadil A. Mohamed (2018). "Link Prediction in Dynamic Social Networks: A Literature Review." IEEE CiSt'18 5th Edition International IEEE Congress on Information Science and Technology, Marrakech, Morocco, October 21 - 27, 2018.

[2] Wang, P., Xu, B., Wu, Y. and Zhou, X. (2014). Link prediction in social networks: the state-of-the-art. *arXiv preprint arXiv:1411.5118*.

[3] Al Hasan, M. and Zaki, M. J. (2011). A survey of link prediction in social networks. In *Social network data analytics* (pp. 243–275). Springer US.

[4] Bliss, C. A., Frank, M. R., Danforth, C. M. and Dodds, P. S. (2014). An evolutionary algorithm approach to link prediction in dynamic social networks. *Journal of Computational Science*, *5*(5), 750–764.

[5] Tylenda, T., Angelova, R. and Bedathur, S. (2009, June). Towards time-aware link prediction in evolving social networks. In *Proceedings of the 3rd workshop on social network mining and analysis* (p. 9). ACM.

[6] Mori J, Kajikawa Y, Kashima H, et al. Machine learning approach for finding business partners and building reciprocal relationships. Expert Systems with Applications, 2012, 39: 10,402–10,407.

[7] Wu S., Sun J., Tang J. Patent partner recommendation in enterprise social networks. In: Proceedings of the 6th ACM International Conference on Web Search and Data Mining (WSDM'13), Rome, Italy, 2013. 43–52.

[8] Berlusconi G., Calderoni F., Parolini N., Verani M., Piccardi C. (2016). "Link Prediction in Criminal Networks: A Tool for Criminal Intelligence Analysis." PLOS ONE, DOI:10.1371/journal.pone.0154244, April 22, 2016.

[9] Almansoori W., Gao S., Jarada T. N., et al. Link prediction and classification in social networks and its application in healthcare and systems biology. Network Modeling Analysis Health Informatics Bioinformatics, 2012, 1: 27–36.

[10] Huang Z., Lin D. K. J. The time-series link prediction problem with applications in communication surveillance. INFORMS Journal on Computing, 2009, 21: 286–303.

[11] Li D., Zhang Y., Xu Z., Chu D. and Li S. (2016). "Exploiting Information Diffusion Feature for Link Prediction in Sina Weibo." Scientific Reports vol(6), article number: 20058(2016).

[12] Liu F., Liu B., Sun C., Liu M., and Wang X. (2015). "Deep Belief Network-Based Approaches for Link Prediction in Signed Social Networks." Entropy, Vol(2015), No(17), PP. 2140-2169; doi:10.3390/e17042140.

[13] Zhang H., Wu1 G. and Ling Q. (2019). "Distributed Stochastic Gradient Descent for Link Prediction in Signed Social Networks." Journal on Advances in Signal Processing, (2019) 2019:3. https://doi.org/10.1186/s13634-019-0601-0.

[14] Bastami E., Mahabadi A., and Taghizadeh E. (2019). "A gravitation-based link prediction approach in social networks." Swarm and Evolutionary Computation, 44 (2019), PP. 176–186.

[15] Mallek S., Boukhris I., Elouedi Z., and Lefèvre E. (2018). "Evidential link prediction in social networks based on structural and social information." Journal of Computational Science, 30 (2019), PP. 98–107.

[16] Ghazaleh Beigi, Suhas Ranganath, and Huan Liu. (2019). "Signed Link Prediction with Sparse Data: The Role of Personality Information." In Companion Proceedings of the 2019 World Wide Web Conference (WWW '19 Companion), May 13–17, 2019, San Francisco, CA, USA. ACM, New York, NY, USA. https://doi.org/10.1145/3308560.3316469.

[17] Shinkuma R., Yuki Sugimoto Y., and Inagaki Y. (2019). "Weighted network graph for interpersonal communication with temporal regularity." Soft Computing, (2019) 23, PP. 3037–3051.

[18] Ai J., Su Z., Li Y. (2019). "Link prediction based on a spatial distribution model with fuzzy link importance." Physica A, 527 (2019) 121155.

[19] Aslan S., Kaya B., and Kaya M. (2019). "Predicting potential links by using strengthened projections in evolving bipartite networks." Physica A, 525 (2019), PP. 998–1011.

[20] Bu Z., Wang Y. , Li H., Jiang J. , Wu Z. , and Cao J. (2019). "Link prediction in temporal networks: Integrating survival analysis and game theory." Information Sciences, 498 (2019), PP. 41–61.

[21] Dharavath R. and Arora N. S. (2019). "Spark's GraphX-based link prediction for social communication using triangle counting." Social Network Analysis and Mining (2019) 9:28.

[22] Eberhard L., Trattner C., and Atzmueller M. (2018). "Predicting trading interactions in an online marketplace through location-based and online social networks." Information Retrieval Journal, (2018), https://doi.org/10.1007/s10791-018-9336-z.

[23] Gao H., Huang J., Cheng Q., Sun H., Wang B., and Li H. (2019). "Link prediction based on linear dynamical response." Physica A, 527 (2019), 121397.

[24] Golzardi E., Sheikhahmadi A., and Abdollahpouri A. (2019). "Detection of trust links on social networks using dynamic features." Physica A, 527 (2019), 121269.

[25] Gu S., Chen L., Li B., Liu W., Chen B. (2019). "Link prediction on signed social networks based on latent space mapping." Applied Intelligence, 2019, https://doi.org/10.1007/s10489-018-1284-1.

[26] Jibouni, A., Lotfi, D., EL Marraki, M., Hammouch, A. (2018). "A novel parameter free approach for link prediction." Proceedings - 2018 International Conference on Wireless Networks and Mobile Communications, WINCOM, Marrakesh; Morocco; 16 October 2018 through 19 October 2018.

[27] Kumar, A., Singh, S.S., Singh, K., Biswas, B. (2019). "Level-2 node clustering coefficient-based link prediction." Applied Intelligence, Volume 49, Issue 7, 15 July 2019, Pages 2762-2779.

[28] Dinh, T.N., Nguyen, D.T., and Thai, M.T. (2012) "Cheap, Easy, and Massively Effective Viral Marketing in Social Networks: Truth or Fiction?", Proceedings of the 23rd ACM Conference on Hypertext and Social Media, June, 2012, pp. 165-174.

[29] Monteserin A. and Armentano M. G. (2019). "Influence me! Predicting links to influential users." Information Retrieval Journal, Volume 22, Issue 1-2, 15 April 2019, Pages 32-54.

[30] Mutinda F. W., Nakashima A., Takeuchi K., and Sasaki Y. (2019). "Time Series Link Prediction Using NMF." 2019 IEEE International Conference on Big Data and Smart Computing, BigComp 2019 - Proceedings1 April 2019, Article number 86795022019 IEEE International Conference on Big Data and Smart Computing, BigComp 2019; Kyoto; Japan; 27 February

2019 through 2 March 2019; Category numberCFP1940X-ART; Code 146803.
[31] Wang Z., Liang J., and Li R. (2018). "A fusion probability matrix factorization framework for link prediction." Knowledge-Based Systems, 159 (2018), PP. 72–85.
[32] Wang Z., Wang Y., Ma J., Li W., Chen N., and Zhu X. (2019). "Link prediction based on weighted synthetical influence of degree and H-index on complex networks." Physica A: Statistical Mechanics and its Applications, Volume 527, 1 August 2019, Article number 121184.
[33] Wu J., Shen J., Zhou B., Zhang X., and Huang B. (2019). "General link prediction with influential node identification." Physica A: Statistical Mechanics and its Applications, Volume 523, 1 June 2019, Pages 996-1007.
[34] Yuan, W., He, K., Guan, D., Zhou, L., Li, C. (2019). "Graph kernel based link prediction for signed social networks." Information Fusion, Volume 46, March 2019, PP. 1-10.
[35] Greene, D., Doyle, D. and Cunningham, P. (2010, August). Tracking the evolution of communities in dynamic social networks. In *Advances in social networks analysis and mining (ASONAM), 2010 International Conference* (pp. 176–183). IEEE.
[36] Yao L., Wang L., Pan L., and Yao K. (2016). "Link Prediction Based on Common-Neighbors for Dynamic Social Network." Procedia Computer Science, Vol (83), 2016, PP. 82 – 89.
[37] Liben-Nowell, D. and Kleinberg, J. (2007). The link-prediction problem for social networks. journal of the Association for Information Science and Technology, 58(7), 1,019–1,031.
[38] Rossetti G., Guidotti R., Miliou I., Pedreschi D., and Giannotti F. (2016). "A supervised approach for intra-/inter-community interaction prediction in dynamic social networks." Soc. Netw. Anal. Min. (2016) 6:86, DOI 10.1007/s13278-016-0397-y.
[39] Moradabadi, B. and Meybodi, M. R. (2018). Link prediction in weighted social networks using learning automata. Engineering Applications of Artificial Intelligence Journal, 2018, 70: 16–24.
[40] Moradabadi, B. and Meybodi, M. R. (2018). Link prediction in stochastic social networks: Learning automata approach. Journal of Computational Science, 2018, 24: 313–328.
[41] Wu, Z., Lin, Y., Zhao, Y., Yan, H. (2018). Improving local clustering based top-L link prediction methods via asymmetric link clustering information. Physica A: Statistical Mechanics and its Applications Journal, 2018, 492: 1,859–1,874.
[42] Fu, C., Zhao, M., Fan, L., Chen, X., Chen, J., Wu, Z., Xia, Y., Xuan, Q. (2018). Link Weight Prediction Using Supervised Learning Methods and Its Application to Yelp Layered Network. IEEE Transactions on Knowledge and Data Engineering Journal, 2018, 99.
[43] Julian, K., & Lu, W. (2016). Application of Machine Learning to Link Prediction.
[44] Mohan, A., Venkatesan, R., Pramod, K. V. (2017). A scalable method for link prediction in large real-world networks. Journal of Parallel and Distributed Computing, 2017, 109: 89–101.
[45] Caiyan, D., Chen, L., Li, B. (2107). Link prediction in complex network based on modularity. Journal of Soft Computing, 2017, 21(15): 4,197–4,214.
[46] Bhat, S. Y. and Abulaish, M. (2015). HOCTracker: Tracking the evolution of hierarchical and overlapping communities in dynamic social networks. *IEEE Transactions on Knowledge and Data Engineering*, *27*(4), 1,019–1,013.
[47] Liben-Nowell, D. and Kleinberg, J. (2007). The link-prediction problem for social networks. *Journal of the Association for Information Science and Technology*, *58*(7), 1,019–1,031.
[48] Gao, S., Denoyer, L. and Gallinari, P. (2011, October). Temporal link prediction by integrating content and structure information. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management* (pp. 1,169–1,174). ACM.
[49] Dunlavy, D. M., Kolda, T. G. and Acar, E. (2011). Temporal link prediction using matrix and tensor factorizations. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, *5*(2), 10.
[50] Ibrahim, N. M. A. and Chen, L. (2015). Link prediction in dynamic social networks by integrating different types of information. *Applied Intelligence*, *42*(4), 738–750.
[51] Niladri S., Saptarshi B. , Sukumar N., Sanasam R. S. (2018). Temporal link prediction in multi-relational network. World Wide Web Journal, 2018, 21(2): 395–419.
[52] Tabourier L., Libert A-S., and Lambiotte R. (2016). "Predicting links in ego-networks using temporal information." EPJ Data Science, (2016) 5:1, DOI 10.1140/epjds/s13688-015-0062-0.
[53] Munasinghe L. and Ichise R. (2012). "Time score: A new feature for link Prediction in Social Networks." IEICE Trans. Inf. & Syst., Vol(E95-D), No. 3, March 2012, PP. 821-828.
[54] Oyama S., Hayashi K., and Kashima H. (2012). "Link Prediction Across Time via Cross-Temporal Locality Preserving Projection." IEICE Trans. Inf. & Syst., Vol(E95-D), No. 11, November 2012, PP. 2664-2674.
[55] Sarkar, P. and Moore, A. W. (2006). Dynamic social network analysis using latent space models. In *Advances in Neural Information Processing Systems* (pp. 1,145–1,152).
[56] Juszczyszyn, K., Musial, K. and Budka, M. (2011, October). Link prediction based on subgraph evolution in dynamic social networks. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on* (pp. 27–34). IEEE.
[57] Zaki *et al.*: Protein complex detection using interaction reliability assessment and weighted clustering coefficient. *BMC Bioinformatics* 2013 14:163.
[58] Chen G. , Xu C., Wang J., Feng J., Feng J. (2020). " Robust non-negative matrix factorization for link prediction in complex networks using manifold regularization and sparse learning." Physica A 539 (2020) 122882.
[59] Sarkar, P., Chakrabarti, D. and Jordan, M. (2012). Nonparametric link prediction in dynamic networks. *arXiv preprint arXiv:1206.6394.*
[60] Lü, L. and Zhou, T. (2011). Link prediction in complex networks: A survey. Physica A: statistical mechanics and its applications, 390(6), 1,150–1,170.
[61] Virinchi S., Mitra P. (2013). "Link Prediction Using Power Law Clique Distribution and Common Edges Distribution." In: Maji P., Ghosh A., Murty M.N., Ghosh K., Pal S.K. (eds) Pattern Recognition and Machine Intelligence. PReMI 2013. Lecture Notes in Computer Science, vol 8251. Springer, Berlin, Heidelberg.
[62] Dong, L, Li Y., Yin H., Le H., and Rui M. (2013). "The Algorithm of Link Prediction on Social Network." Mathematical Problems in Engineering, Vol(2013), Article ID 125123, 7 pages, http://dx.doi.org/10.1155/2013/125123.
[63] Zhang Y., Shen S., and Wu Z. (2018). "Improve Link Prediction Accuracy with Node Attribute Similarities." International Conference on Computer Engineering and Networks CENet2018 2018: The 8th International Conference on Computer Engineering and Networks (CENet2018), PP. 376-384.
[64] M. T. Riaz, Y. Fan, J. Ahmad, M. A. Khan, and E. M. Ahmed, "Research on the Protection of Hybrid HVDC System," in 2018 International Conference on Power Generation Systems and Renewable Energy Technologies (PGSRET), 2018, pp. 1–6.
[65] Li, X., Du, N., Li, H., Li, K., Gao, J. and Zhang, A. (2014, April). A deep learning approach to link prediction in dynamic networks. In *Proceedings of the 2014 SIAM International Conference on Data Mining* (pp. 289–297). Society for Industrial and Applied Mathematics.
[66] Pan L., Zhou T., Lü L., and Hu C-K. (2016). "Predicting missing links and identifying spurious links via likelihood analysis." Scientific Reports, Vol(6:22955), DOI: 10.1038/srep22955.
[67] Wu Z., Lin Y., Wan H. and Jamil W. (2016). "Predicting top-L missing links with node and link clustering information in large-scale networks." Journal of Statistical Mechanics: Theory and Experiment, (2016) 083202.
[68] Bhawsar Y. and Thakur G.S. (2016). "Performance Evaluation of Link Prediction Techniques Based on Fuzzy Soft Set and Markov Model.", Fuzzy Information and Engineering, 8:1, 113-126, DOI: 10.1016/j.fiae.2016.03.007.
[69] Menon A.K., Elkan C. (2011) Link Prediction via Matrix Factorization. In: Gunopulos D., Hofmann T., Malerba D., Vazirgiannis M. (eds) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2011. Lecture Notes in Computer Science, vol 6912. Springer, Berlin, Heidelberg.
[70] Han J., M. Kamber, and J. Pei. 2012. Data Mining Concepts and Techniques. Morgan Kaufmann: Waltham, MA 02451, USA, 2012.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 255-262 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Optimization of the Electrical Discharge Machining of Powdered Metallurgical High-Speed Steel Alloy using Genetic Algorithms

Mohd Razif Idris[1], Imad Mokhtar Mosrati[2,*]

[1]*Malaysia Italy Design Institute (MIDI), Universiti Kuala Lumpur, 56100, Malaysia.*

[2] *Mechanical Engineering Department, Faculty of Engineering, Azzaytuna University, 21218, Libya.*

ARTICLE INFO

Article history:
Received: 04 October, 2019
Accepted: 10 November, 2019
Online: 05 December, 2019

Keywords:
EDM machine
Genetic algorithm optimization
Tool wear rate
Micro-crack size
Silver-Tungsten electrode

ABSTRACT

Through the Electrical Discharge Machining, the temperature is very high, which can lead to the material phase's transformation and affects material properties, which can lead to failure of the products in the industry. This study aims to investigate the effect of a new input parameter (pulse cycle time Tc), with other parameters on the EDM responses, prediction models are developed. 3D Laser Microscope is used to measure the surface defects. And to assess the effect of using a silver-tungsten electrode on improving EDM performance during machining of powdered metallurgical high-speed steel which is a very important material in the industry, such as aerospace, automotive, and electrical power industries, it is used in structural components and heat exchangers. An integrated approach combines Response surface design and Genetic Algorithm optimization methodology for modeling and optimizing the process responses is utilized. Two software's Minitab software and Matlab software were used for this purpose. The experimental results demonstrated comparative findings when the silver-tungsten electrode was used, and the optimization percentage was increased when the Genetic Algorithms method was used. Optimized values of the responses were as following: Micro-crack size was 0.965 µm, the optimization percentage is 20.64%. Tool Wear Rate was 0.0020 g/min, the optimization percentage is 13.04 %.

1. Introduction

The product quality relies on the manufacturing process, optimization tools play a key role in enhancing product quality. Therefore the interest of this study to optimize the parameters and responses during the Electrical Discharge Machining EDM process of the material Powdered metallurgical high-speed steel. EDM is a nonconventional manufacturing process, utilized for very high hardness material, which can not be machined with traditional processes like turning or milling. It is a thermoelectrical process in which there is spark generation between the electrode and the workpiece. Response surface design methodology (Box Behnken) and genetic algorithm (GA) were integrated to optimize the input parameters and process responses. Genetic Algorithm (GA) Optimization Methodology is an artificial intelligence prediction tool, it has been getting interested to solve problems that are hardly solved by the use of conventional methods.

A review of the literature undertaken found that most researches likely used the same ideas for optimization of process responses, like Taguchi, and study the same input parameters. Here is a review of the literature on the EDM process optimization based on the publishing year of the researches.In [1], the researcher improved process performance such as metal removal rate, surface roughness and overcut based on Grey Taguchi technique for EDMing of Ti-6A-4V Titanium alloy. The process parameters selected in this study are current, T-on, voltage, T-off, and duty factor. The optimized process parameters result in a lower wear ratio of the electrode, higher MRR, and better SR, the study shows an improved TWR of 15%, the MRR of 12% and SR of 19%. In [2], the authors established an analytical study for investigating the impact of input parameters on the MRR, SR, and RLT. The copper electrode is utilized, the other inputs are current and T-on. The best selection of input parameters and EDM

* Imad Mokhtar M., Email: imadmokhtar@yahoo.com

machining conditions is performed. Study [3] used the Taguchi design methodology and Grey Relational Analysis tool to improve the outputs like MRR, TWR, and SR of mild steel IS 2026 during the EDM process. It is found the optimal pulse current at 26A, pulse on time at 55µs and pulses off time at 5µs.

The impact of operating parameters like voltage, pulse on-time, pulse of-time and current on outputs such as MRR and SR of AISI D2 workpiece EDM, a copper electrode is investigated [4], the experiment design was made using L9 OA and using response surface methodology to improve the outputs, it was discovered that the current is the most significant parameter influencing the MRR and the SR. Another study [5], authors researched the impact of wire EDM on aerospace material fracture toughness, Grade 5 titanium (Ti6AI4V), fracture toughness is widely used for damage tolerance analysis of aerospace components in which critical crack sizes are computed for given loading conditions to arrive at safe inspection and maintenance intervals. The compact tension was used to evaluate the fracture toughness. Obtained results indicate a slight decrease in fracture toughness compared to those reported in the literature.

In [6] the authors studied electrical discharge machining (EDM) of insulating ceramic materials, such as diamond, using the assisting electrode method. This study investigated the dependence of the machining properties on amounts of additives, three types of insulating Si3N4 ceramics were machined by sinking EDM and by wire EDM. The MRR and the SR were estimated for the EDMed material. Different dependencies on the amounts of added Al2O3 were observed, and the MRR depended on the size of the electrode. In [7], the researcher introduced a micro-EDM electrothermal system to simulate the crater forming. This model integrates machining conditions such as heat flux, temperature-dependent thermal properties, and plasma radius, the heat transfer equation and its calculations of the produced crater are used to assess the fraction of energy distribution to the electrodes, results show a good agreement with experimental results. The parameters that affect the performance output (MRR) of the workpiece, parameters were chosen for optimization are a T-on, T-off followed by the current to examine the impact of the process inputs on the MRR. For AISI D2 die steel, brass is the used electrode material [8].

In [9], authors implemented an approach to assess the MRR in the EDM process. In this study, a finite element model was used to simulate and investigate the process parameters and their impact on outputs such as MRR of die steel. The optimal process inputs combination at the machining of alumina using the electrochemical machining process is investigated [10], the main input parameters are electrolyte concentration, voltage and feed rate, machining performance are MRR and TWR, graphite was used as electrode material.

In [11], authors presented a study to investigate the overcut and taper angles in the machining of stainless steel 304, parameters were current, T-on and T-off. The experimentation is performed using the RSM method, the current was found to be the most significant factor affecting the overcut and the taper angle followed by T-on and T-off, as the current and T-on increase, the overcut and taper angle is increased, so as T-off increases, the two responses decrease, the optimized parameters that give the minimum overcut and taper angle response are 9 A current, 30 µs Pulse on time and 35 µs Pulse off time which gives 0.010 mm overcut and 0.244° taper angle. Grey relational analysis for Dry Electrical Discharge Machining is used in [12], to investigate and optimize the process parameters T-on, T-off, current, and voltage on the SR using the deionized water and air as a dielectric liquid for workpiece EN-8. ANOVA was performed to find the significant parameters on responses, found that current and T-on were the most significant input parameters.

It is clear from previous studies, manufacturers and researches relied on the use of non-traditional manufacturing methods like (EDM) instead of traditional methods like turning and milling when manufacturing the metals and alloys with high hardness such as the study task material, that could not be machined in traditional ways because the hardness of the workpiece is higher than the hardness and rigidity of the machine and cutting tool.

The main contribution of this study introducing an integrated approach combines both the Box-Behnken design and modeling method and the Genetic Algorithm as an optimization methodology that can be an effective design and optimization tool. And to investigate the influence of electrode materials (Silver-Tungsten) and pulse cycle time as a new input parameter on the EDM responses during machining of powdered metallurgical high-speed steel alloy.

2. Materials, Experimental Procedure and Data Setup

This part consists of the experimental setup, selection of workpiece, equipment and tool materials, design and optimization experimental methodologies.

2.1 Machine

All experiments were carried out on the EDM machine model (CHARMILLES Roboform 40). The machine is complete with programming. It features a PC-based CNC control with an LCD screen for faster data processing and servo control. Figure 1 shows the used EDM machine.



Figure 1: EDM machine

2.2 Workpiece Material

The task material in this study is Powdered Metallurgical High-Speed Steel Alloy, the nearest equivalent material is (AISI

M3-2). It is a very important material in the industry; it has not been investigated in the previous literature. Table 1 displays chemical compositions of the material [16], figure 3 shows a sample of material applications.

The chemical compositions as following:

Table 1: Chemical composition of the workpiece material

C	Cr	Mo	W	V
1.30	4.00	5.00	6.50	3.00

Material Properties: Hardness (187.6 HV), Density (0.008166 g/mm3). The shape of the workpiece is a plate, its width 40 mm, length of 60 mm and depth 10mm, required to machine a hole with 20mm diameter, and with 2mm depth. Figure 2 shows the shape and dimensions of the workpiece.

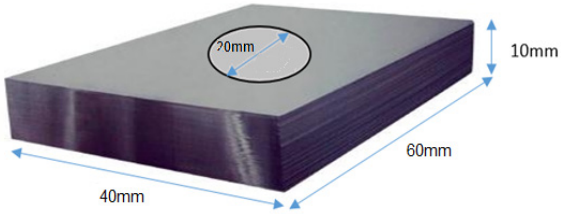


Figure 2: Dimensions of workpiece Material

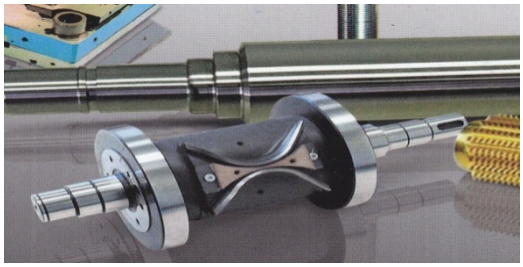


Figure 3: Sample of the workpiece material

2.3 Electrodes

The electrode material used is silver-tungsten, had 100 mm long, diameter 20.70mm. Silver tungsten is strong, capable of penetrating the toughest materials, and machinable. Silver has great conductivity at high temperatures and pressures. Figure 4 shows the electrode shape. Table 2 shows the Physical and Mechanical Properties of Silver - Tungsten electrode [17].



Figure 4: Shape of the electrodes

Table 2: Physical and mechanical properties of silver-tungsten electrode material

Nominal Composition (% Weight)	Density (g/cm^3)	Electrical Resistivity (20°C, μΩ•cm) ≤	Electrical Conductivity (IACS%, ≥)	Hardness (HB, ≥) Kgf/mm^2	Bending Strength (MPa)
25 Silver 75Tungsten	15.40	4.2	41	165	686

2.4 The 3D Laser Microscope Instrument

The 3D Laser Scanning Microscope was utilized for measuring Micro-Cracks size and taking the topography images. The magnification used in this inspection was 1000X. Figure 5 shows the 3D Measuring Laser Microscope. Due to huge temperature in the EDM process, and rapid heating and quenching, surface defects are created on the workpiece such as cracks, recast layer, surface roughness, and craters, in this study Micro- Cracks Size will be investigated. 3D Laser Microscope was used to measure the surface response in units of a micrometer, with a magnification of 1000.

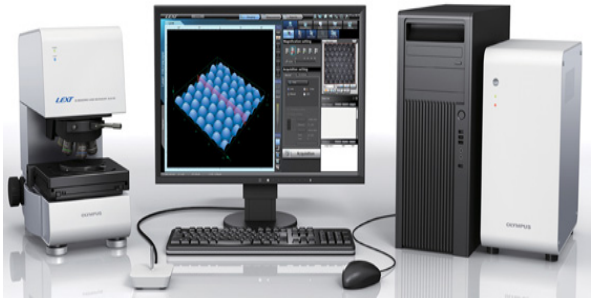


Figure 5: The 3D Laser Microscope device

2.5 EDM Process Responses and surface responses

Industries like automotive and aerospace have used the EDM process to produce products. "The aerospace industry used the EDM process to manufacture some aircraft parts because of the intricate shapes, tough alloys, and very tight tolerances involved, but it also recognized the dangers of the damaged surfaces and product defects resulting from the process" [13]. The process responses will be investigated and optimize in this research were: Tool Wear Rate and Micro-Crack Size.

Tool Wear Rate (TWR):
The wear rate of the electrode was determined using the formula:

$$\text{TWR (g /min)} = \frac{Eb - Ea\ (\text{g})}{\text{MT (min)}} \quad (1)$$

Where (Eb) and (Ea) are the weights of electrode material before and after machining, respectively, and MT is the machining time.

2.6 Experimental Setup

In this study Box-Behnken Experimental design Methodology was used to design the EDM process for Powdered

Metallurgical High-Speed Steel Alloy, L30 Orthogonal Array (OA) is designed, with (3) center points. Machining parameters are three continuous factors: Pulses cycle time (Tc), also it is called (U) as shown in Figure 6 [14]. Dielectric Fluid Flushing Pressure (P), and Voltage (V), one category factor: (the type of electrode material). Table 3 displays the input parameters and their values.

Fixed parameters: current 6A, Pulse width 200 µs, the polarity of the electrode is positive. Dielectric hydrocarbon fluid with a density of 0,783 g / ml and a flashpoint of 125 ºC is used. Flushing method: the dielectric fluid was injected through side pipe novels providing adequate flushing to remove the eroded particles away from the working gap.

Table 3*:* Machining parameters and their levels

Parameters	Symbol	Levels			Unit
		1	2	3	
Period of machining (Pulses cycle time)	Tc	0.1	6.4	12.8	Sec
Dielectric fluid Flushing pressure	P	5	15	25	Psi
Machining voltage	V	120	160	200	V
Electrode material	Ag-W Silver-Tungsten				

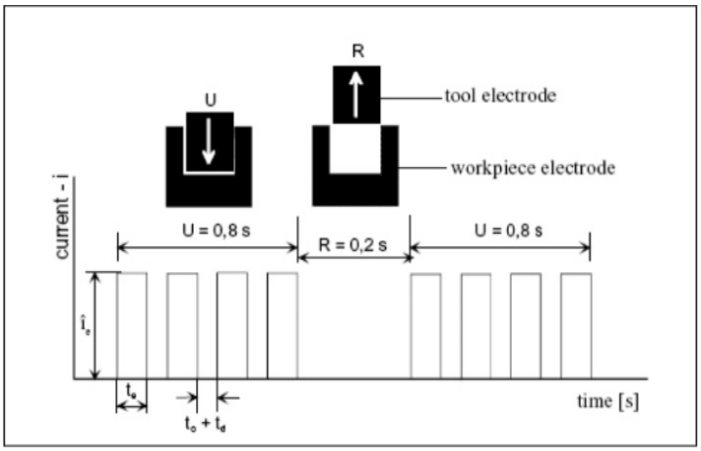


Figure 6*:* Series of pulses cycle time Tc, (U)

2.7 Genetic Algorithms Optimization Tool (GA)

"Genetic Algorithms, GA contains a chromosome, set of population, fitness function, mutation and selection, it begins with a set of solutions represented by chromosomes, called population, solutions from one population are taken and used to form a new population, which is motivated by the possibility that the new population will be better than the old one, further, solutions are selected according to their fitness to form new solutions, this process is repeated until some condition is satisfied. Algorithmically, the basic of GA is contains the following actions, start, fitness, new population, replace, test, and loop " [15].

Figure 7 shows the flowchart of the Genetic Algorithm optimization methodology.

In this research the models developed by (Box – Behnken) design methodology will be used as an objective function, the upper and lower bounds of parameters are identified in the GA optimization technique as below. The aim was to minimize the Response value.

Minimize f(x) = Response

Responses to be as minimum as better

Subjected to,

$120 \leq V \leq 200$;

$0.1 \leq Tc \leq 12.8$;

$5 \leq P \leq 25$.

GA Toolbox of MATLAB software is used. All the previous steps of the Genetic Algorithm will be done using Matlab software, after achieving and building the models of all the responses from the (Box – Behnken) optimization methodology to be used as objective function equations, all results obtained from the runs are presented in the following chapters.

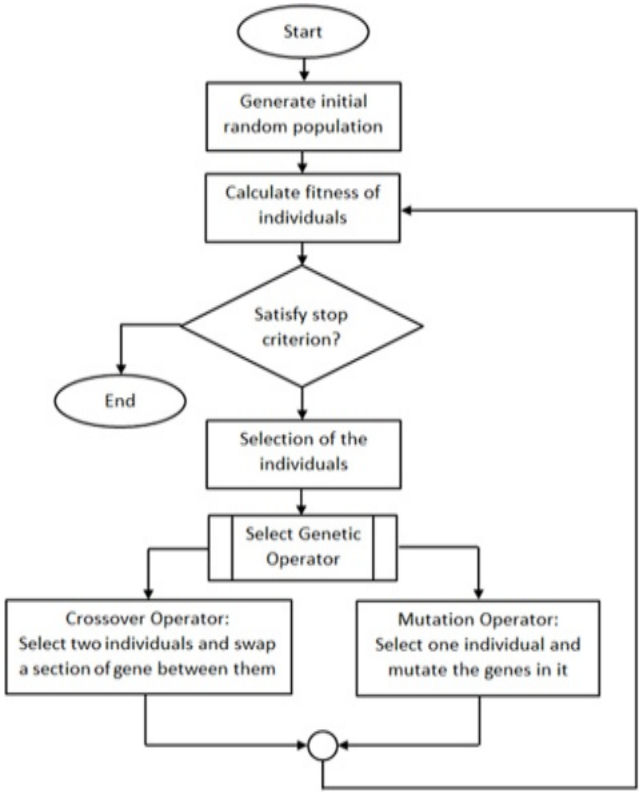


Figure 7: Flowchart diagram of Genetic Algorithm optimization methodology

3. Results and Analysis

In this section, the results obtained from the running of Minitab design software and Matlab GA optimization software are discussed for each response variable, tool wear, micro crack size. Then a regression analysis was conducted to find the best-fit models. The main effects and interaction effects were discussed. Table 4 shows the Experimental results.

Table 4: Experimental Results

Run Order	V Volt	Tc Sec	P	TWR g/min	Micro-crack size µm
1	160	0.10	5	0.0059	1.344
2	120	6.45	25	0.0071	2.363
3	120	6.45	5	0.0088	1.688
4	160	6.45	15	0.0091	1.905
5	160	12.8	25	0.0142	2.026
6	160	12.8	5	0.0087	1.608
7	160	6.45	15	0.0090	1.817
8	160	0.10	25	0.0023	1.350
9	120	12.8	15	0.0093	1.721
10	200	6.45	5	0.0083	2.001
11	200	12.8	15	0.0175	2.157
12	120	0.10	15	0.0033	1.905
13	200	0.10	15	0.0047	1.216
14	160	6.45	15	0.0085	1.616
15	200	6.45	25	0.0122	1.794

The Minitab17 software package was used to estimate the models of the data obtained from the experiments. Figures 8-23 Show the 2D plots, to analyze the main effect and the mutual interactions between the input machining parameters (V, Tc, P, and electrode material) on the responses. The ANOVA analysis is shown in tables 5 - 6. In this analysis, all parameters with a p-value of less than 0.05 are significant.

Table 5*:* Analysis of Variance of TWR

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	6	0.000208	0.000035	29.40	0.000
Linear	3	0.000168	0.000056	47.44	0.000
V	1	0.000025	0.000025	21.41	0.002
Tc	1	0.000140	0.000140	119.14	0.000
P	1	0.000002	0.000002	1.78	0.218
2-Way Interaction	3	0.000040	0.000013	11.35	0.003
V*Tc	1	0.000012	0.000012	9.82	0.014
V*P	1	0.000008	0.000008	6.66	0.033
Tc*P	1	0.000021	0.000021	17.58	0.003
Error	8	0.000009	0.000001		
Lack-of-Fit	6	0.000009	0.000002	14.86	0.064
Pure Error	2	0.000000	0.000000		
Total	14	0.000217			

S	R-sq	R-sq(adj)	R-sq(pred)
0.0010851	95.66%	92.41%	77.72%

3.1 Effect of Process Parameters on Tool Wear Rate TWR

Statistical inferences:

1. The model is significant, as seen p-value is 0.000, that is meant the model is significant and adequate.
2. Lack of Fit is 0.064, it is not significant.
3. The "R-Squared" of 95.66%, "Pred R-Squared" of 77.72 %, the "Adj R-Squared" of 92.41%, indicates an adequate signal.

4. Values of "Prob. less than 0.05" indicate model terms are significant. In this case, V, Tc, (Tc*P), (V*P), and (V*Tc) are significant model terms.

Equation model of TWR:

$$TWR = 0.01524 - 0.000051\,V - 0.000949\,Tc - 0.000740\,P + 0.000007\,V*Tc + 0.000004\,V*P + 0.000036\,Tc*P$$

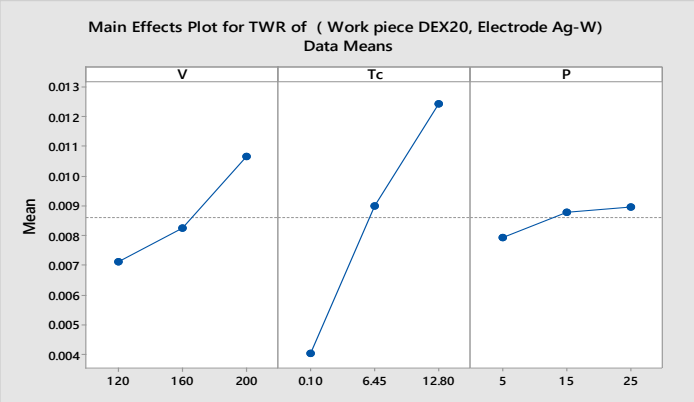


Figure 8: Main effects plot for TWR

The main effects: Figure 8 displays the main effects plot for TWR. As the voltage increased as TWR increased. Increasing pulse cycle time resulted in increasing of TWR, the minimum value in TWR was at 0.10s. Dielectric liquid pressure is increased as TWR stays almost no change, no significant effect.

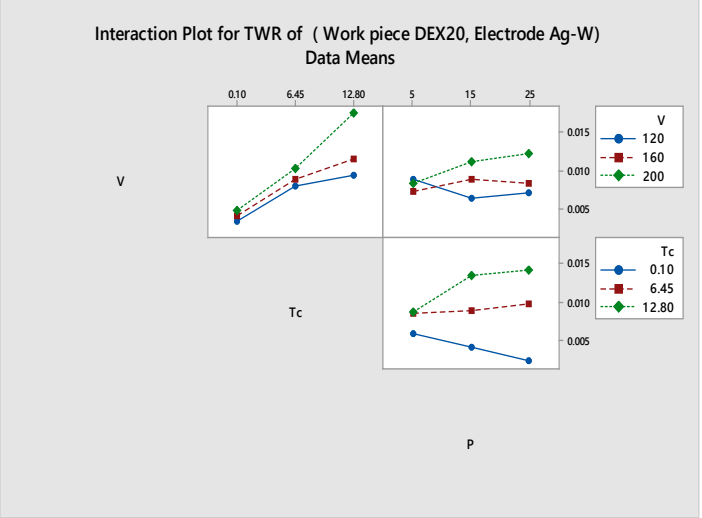


Figure 9*:* Interaction effects plot for TWR

The interaction effect: Figure 9 displays the interaction effects plots for TWR. Based on the ANOVA table, the p-value of all interactions (Tc*P), (V*P), (V*Tc) was less than 0.05. Therefore, the interaction effects were significant.

3.2 Effect of process parameters on Micro-crack Size

Statistical inferences:

1. P-value of the model is 0.004, so it is significant and adequate.
2. The "R-Squared" of 90.22%, "Pred R-Squared" of 87.74 %, the "Adj R-Squared" of 96.51%. This model can be used to navigate the design space and indicates an adequate signal.
3. Tc, P, (V*V), (Tc*Tc), (V*Tc), and (V*P) are significant terms.

Regression Equation of Micro-C racks:

M-Crack Size = 4.232 - 0.0355 V - 0.1027 Tc + 0.0867 P + 0.000109 V*V- 0.00507 Tc*Tc+ 0.000072 P*P + 0.001107 V*Tc - 0.000551 V*P + 0.001622 Tc*P

Table 6*:* Analysis of variance for Micro-crack size

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	9	1.33501	0.148334	15.35	0.004
Linear	3	0.49182	0.163940	16.96	0.005
V	1	0.03239	0.032385	3.35	0.127
Tc	1	0.35998	0.359976	37.25	0.002
P	1	0.09946	0.099458	10.29	0.024
Square	3	0.28987	0.096622	10.00	0.015
V*V	1	0.11302	0.113023	11.70	0.019
Tc*Tc	1	0.15448	0.154476	15.99	0.010
P*P	1	0.00019	0.000192	0.02	0.893
2-Way Interaction	3	0.55332	0.184441	19.09	0.004
V*Tc	1	0.31641	0.316406	32.74	0.002
V*P	1	0.19448	0.194481	20.13	0.006
Tc*P	1	0.04244	0.042436	4.39	0.090
Error	5	0.04832	0.009664		
Lack-of-Fit	3	0.00443	0.001476	0.07	0.972
Pure Error	2	0.04389	0.021944		
Total	14	1.38333			

S	R-sq	R-sq(adj)	R-sq(pred)
0.0983035	96.51%	90.22%	87.74%

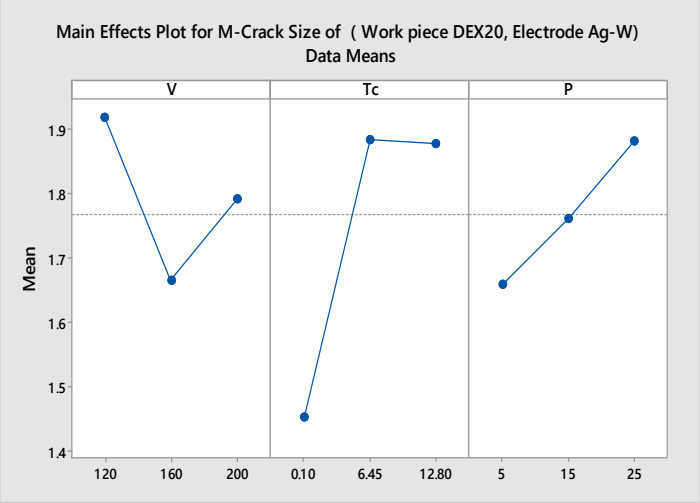


Figure 10: Main effects plot for Micro-Crack Size

The main effects: based on the plot in figure 10, the Micro-crack size decreased when the voltage increased until 160v then it increases. And at the pulse cycle time increasing resulted in increasing of Micro-crack size until 6.45s then it decreases slightly. Dielectric liquid pressure increased as the Micro-crack size increased. The minimum value was at P 5psi, Tc 12.8s, and V 160v.

The significant interaction effect was only between voltage and pulse cycle time (V*Tc), and between voltage and liquid pressure (V*P), from the analysis table, the p-value of this interaction effect was less than 0.05 Therefore, this interaction effect was significant.

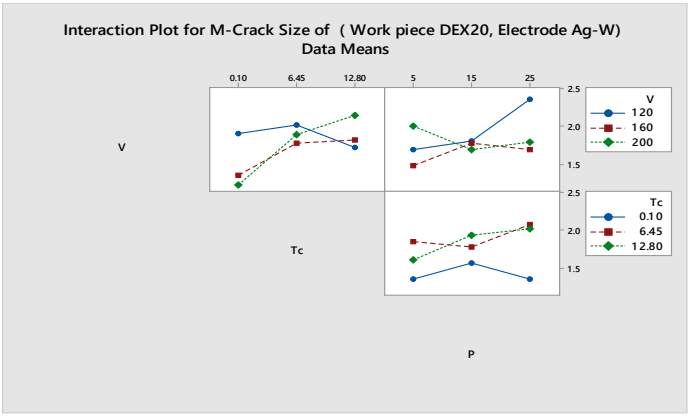


Figure 11: Interaction effects plot for Micro-Crack Size,

3.3 Topography Images using 3D Laser Microscope

The following image depicted how surface responses changed by the process parameters and electrode types. Figure 12 shows the best results for the surface defects got from performing experiments. The micrograph of EDMed surfaces with the high magnification (1000×), for the machining parameter combination V/Tc/P of 200v/0.1s/15psi is shown in figure 12. Theses machining combinations produced minimum Micro-crack width 1.216µm.

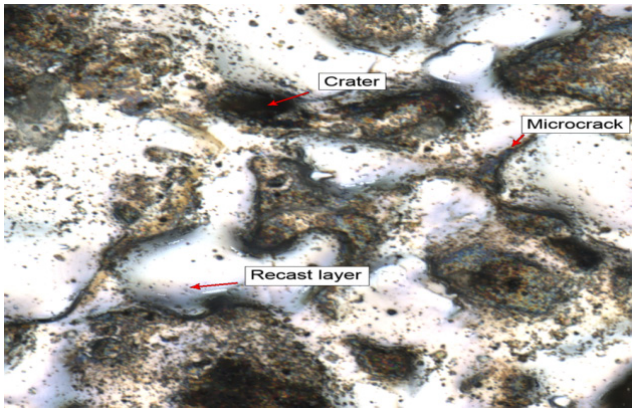


Figure 12: Minimum Micro-crack size 1.216µm, at V =*200v*, Tc=0.1Sec, P= 15psi

4. Responses Optimization using Genetic Algorithm methodology

The equations developed from the Response surface design in the last section will be used as objective functions in the Genetic Algorithm optimization, the target of these responses will be minimized.

Subjected to,
$120 \leq V \leq 200$;
$0.1 \leq Tc \leq 12.8$;
$5 \leq P \leq 25$;
Tool Material = Ag-W.
The equation variables and the input parameters are changed to the coding system that accepted in the Genetic Algorithm Tool.
Response = f (x)

V= X1, Tc=X2, P=X3, Responses=f(x), Maximum and minimum boundary conditions are [120, 0.1, 5] and [200, 12.8, 25]. Tables 7-8 and Figures 13-14 show the optimization results.

4.1 Optimization of Tool Wear Rate TWR using Genetic Algotithm

Response surface equation:

TWR = 0.01524 - 0.000051 V - 0.000949 Tc - 0.000740 P + 0.000007 V*Tc + 0.000004 V*P + 0.000036 Tc*P

Coding equation (minimization):

f(x) = 0.01524 - 0.000051*x(1) - 0.000949*x(2) - 0.000740*x(3) + 0.000007*(x(1)*x(2))+ 0.000004*(x(1)*x(3)) + 0.000036*(x(2)*x(3));

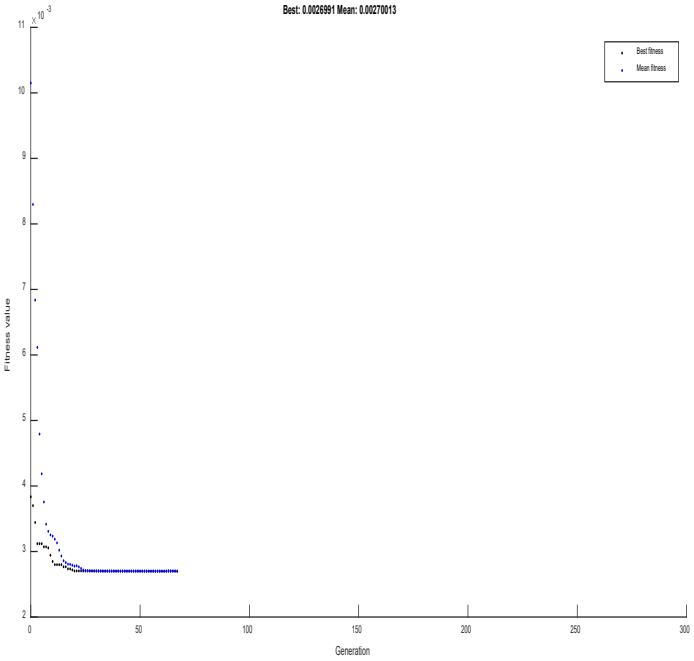


Figure 13: Plots of fitness value with the number of generations of TWR

Table 7: The optimum values of process parameters and optimized value of Tool Wear Rate

Response	Optimize the value of input parameters			Value g/mm	Iteration Number
	V	Tc	P		
TWR	120	0.1	25	Objective function value: 0.002699	67

4.2 Optimization of Micro-Crack Size using Genetic Algotithm

Response surface equation:

M-Crack Size = 4.232 - 0.0355 V - 0.1027 Tc + 0.0867 P + 0.000109 V*V- 0.00507 Tc*Tc+ 0.000072 P*P + 0.001107 V*Tc - 0.000551 V*P + 0.001622 Tc*P

Coding equation (minimization):

f(x) = 4.232 - 0.0355*x(1) - 0.1027*x(2) + 0.0867*x(3) + 0.000109*(x(1)^2) -0.00507*(x(2)^2)+ 0.000072*(x(3)^2) + 0.001107*(x(1)*x(2)) - 0.000551*(x(1)*x(3)) + 0.001622* (x(2)*x(3));

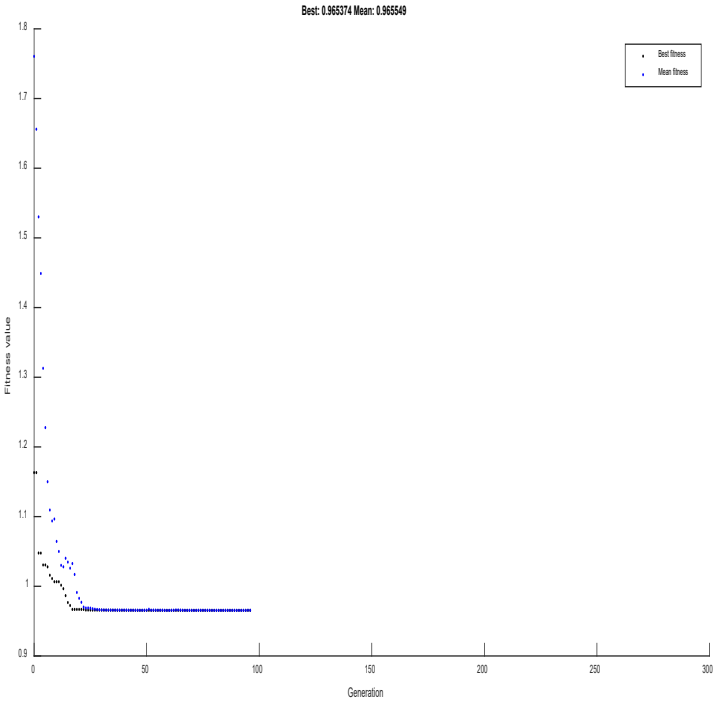


Figure 14: Plots of fitness value with the number of generations of Micro-Crack Size

Table 8: The optimum values of process parameters and the optimized value of Micro-Crack

Response	Optimize the value of input parameters			Value µm	Iteration Number
	V	Tc	P		
M-Crack Width	200	0.1	25	Objective function value: 0.965	96

5. Conclusion

This research demonstrated very good results of the modeling and optimization of TWR, and Micro Crack size for the material Powdered Metallurgical High-Speed Steel Alloy during the EDM process using Box Behnken Design Method with Genetic Algorithm optimization methodology to increase the optimization percentage. The results of using Silver – Tungsten electrode were very good and showed that it has an important impact on the output values, this leads to thinking about not relying on the use of copper electrode only as is clear from the previous studies. This research contributed to highlighting the importance of using genetic algorithms as a method to improve the outputs of the EDM process. Also contributed to study and investigate a new input parameter called Pulses Cycles Time Tc, that it has not been studied previously, results demonstrated it has an effective effect on improving the outputs of the manufacturing process

The summary of optimal results of TWR and Micro Crack Size for EDMing of the material Powdered Metallurgical High-Speed Steel Alloy using Silver-Tungsten electrode to be as follows:

- For Tool Wear Rate TWR: Tc and V are significant parameters. P is not significant. The model is significant, where is R2, 95.66% and Lack of Fit 0.064**.** Minimum Value of TWR as predicted from experiments was 0.0023 g/min. Optimized value using Genetic Algorithm is 0.0020 g/min, if the applied machining conditions are: V=120v, Tc=0.1s, P= 25psi, optimization percentage using GA is 13.04 %.

- For Micro-crack size: Tc is the most significant parameter by then P. V is not significant, the model is significant where R^2 96.51% and Lack of Fit 0.972. The minimum Value of Micro-crack size as predicted from experiments was 1.216 µm. Optimized value using Genetic Algorithms was 0.965 µm when the applied machining conditions are: V=200v, Tc=0.1s, P= 25psi, optimization percentage using GA is 20.64%.

These findings will enable EDM researchers and manufacturing engineers to select the correct parameters for Powdered Metallurgical High-Speed Steel material EDMing.
The models and the data in this work are expected to become a powerful tool in the hands of EDM researchers to extend the system understanding, and also to improve the process performance.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] J. Y. Kao, C. C. Tsao, S. S. Wang, and C. Y. Hsu. Optimization of the EDM parameters on machining Ti-6Al-4V with multiple quality characteristics. Int. J. Adv. Manuf. Technol, 2010.

[2] M. Gostimirovic, P. Kovac, M. Sekulic, and B. Skoric. Influence of discharge energy on machining characteristics in EDM. J. Mech. Sci. Technol, 2012.

[3] S. Raghuraman, K. Thiruppathi, T. Panneerselvam, and S. Santosh. Optimization of Edm Parameters Using Taguchi Method and Grey Relational Analysis for Mild Steel Is 2026. Int. J. Innov. Res. Sci. Eng. Technol, 2(7): 3095–3104, 2013.

[4] S. B. Chikalthankar, V. M. Nandedkar, and S. V Borde. Experimental Investigations of EDM Parameters, 7(5): 31–34, 2013.

[5] D. M. Madyira and E. T. Akinlabi. Effects of Wire Electrical Discharge Machining on Fracture Toughness of Grade 5 Titanium Alloy. Int. J. Mech. Prod. Engineering, 2: 1–5, 2014.

[6] D. Hanaoka, Y. Fukuzawa, Y. Kaneko, and T. Harada. Discharge Machining of Insulating Si3N4 Ceramics with added Al2O3. Int. J. Electr. Mach. 19: 45–50, 2014.

[7] B. Shao and K. P. Rajurkar. Modeling of the crater formation in micro-EDM. In Procedia CIRP, 2015.

[8] J. Jeykrishnan, B. Vijaya Ramnath, A. Jude Felix, C. Rupan Pernesh, and S. Kalaiyarasan. Parameter optimization of electro-discharge machining (EDM) in AISI d2 die steel using Taguchi technique. Indian J. Sci. Technol, 9(4), 2016.

[9] A. Gosavi and A. B. Gaikwad. Predicting Optimized EDM Machining Parameter through Thermo Mechanical Analysis. IOSR J. Mech. Civ. Eng., 13(3): 2320–334, 2016.

[10] M. Goud, A. K. Sharma, and S. Zafar. A Taguchi approach to optimize electrochemical discharge machining of alumina, 2016.

[11] E. A. H. Hanash. Investigation of the Effect of Edm Parameters on Dimensional Accuracy of Drilled Holes of 304 Stainless Steel Using Rsm Method By International Islamic University. Int. Islam. Univ. MALAYSIA, 2017.

[12] N. Singh Khundrakpam, G. Singh Brar, and D. Deepak. Grey-Taguchi Optimization of Near Dry EDM Process Parameters on the Surface Roughness. In Materials Today: Proceedings, 5(2): 4445–4451, 2018.

[13] Aerospace Manufacturing, EDM for Aerospace, Aerospace Production, Radical Departures. Retrieved from https://www.radical-departures.net/articles/advances-in-edm-for-aerospace, 2019.

[14] I. A. Amorim, F.L.; Schäfer, G.; Stedile, L.J.; Bassani. On the Behavior of Parameters and Copper -Tungsten Electrode Edge Radius Wear When Finishing Sinking EDM of Tool, 35, 2010.

[15] M. N. Ab Wahab, S. Nefti-Meziani, and A. Atyabi. A comprehensive review of swarm optimization algorithms. PLoS One, 10(5): 1–36, 2015.

[16] Special steel and Alliance. Retrieved from http://www.specialsteel.com.my/page.php?page=powder, (2019)

[17] Stanford Materials Corporation. Retrieved from: http://www.stanfordmaterials.com/Tungsten-copper.html, (2019)

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 256-262 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Hospital Warehouse Management during the construction of a new building through Lean Techniques

Carlotta Patrone*,[1], Maryna Mezzano Kozlova[2], Monica Brenta[3], Francesca Filauro[4], Donatella Campanella[4], Antonietta Ribatti[5], Elisabetta Scuderi[5], Tiziana Marini[6], Gabriele Galli[2], Roberto Revetria[2]

[1]*E.O. Ospedali Galliera, General Directorate, Mura delle Cappuccine 14, 16128, Genoa, Italy*

[2]*University of Genoa, Mechanical, Energetic, Industrial and Transport Engineer Department (D.I.M.E.), Via Montallegro 1,16145 Genoa, Italy*

[3]*E.O. Ospedali Galliera, Anaesthesia Department, Mura delle Cappuccine 14, 16128, Genoa, Italy*

[4] *E.O. Ospedali Galliera, Pharmacy, Mura delle Cappuccine 14, 16128, Genoa, Italy*

[5] *E.O. Ospedali Galliera, Warehouse and Supply Department, Mura delle Cappuccine 14, 16128, Genoa, Italy*

[6] *E.O. Ospedali Galliera, Information and Technology Department, Mura delle Cappuccine 14, 16128, Genoa, Italy*

ARTICLE INFO

Article history:
Received: 22 October, 2019
Accepted: 23 December, 2019
Online: 27 January, 2020

Keywords:
Business Process Reengineering (BPR)
Lean
Management
Warehouse
Supply Chain Resiliency
Design of Experiments
Healthcare

ABSTRACT

This paper shows a method applied during the reconstruction of an existing middle size Italian hospital. In fact, during this delicate phase, the project had a significant decrease in storage space and access gates for goods, medicines and medical devices. This aspect has a negative impact on road traffic. There was no consensus among the management on how to deal with this problem. In addition, a great tension between engineers and warehouse managers made it impossible to find a solution. For the best of our knowledge there is no available literature on this topic. Hence, a deep analysis had been carried out according to the Business Process Reengineering (BPR). For this reason, a hospital lean group has been created and a statistically study on storage volume inside the warehouse and a monitoring of trucks and goods arrival has been implemented. Three different solutions had been identified through data collection, design of experiments and lean application. The first one was focused on an exclusively internal management, the second one on an exclusively external management whereas the last one was focused on a partially internal scenario. Advantages, disadvantages, costs and the impact on the road traffic has been highlighted for each scenario. Afterward the Management had all the elements to make the best decision. The method reported in this study could be extended during the construction or reconstruction of any company that needs a warehouse or to face a problem in a tense atmosphere.

1. Introduction

Warehouse management inside hospitals has always been considered a crucial topic. Venkateswaran highlighted the complexity of the management of drugs warehouses for the uncertain demand and the high risk of product unavailability that could potentially compromise the safety of the patients [1]. One of the more critical issues is the difficulty to apply Lean Technologies inside the warehouse. According to Frank, warehouses and Lean are considered as mutually exclusive due to the strong difficulties of application of this method in a stockage setting. Nonetheless, the potential advantages are to reduce the "non-value-added steps" as much as possible and to optimize the stock flow [2].

This paper has been developed during the construction of a new building and the main target subjects were the office supply, the medical devices and the drugstore warehouse of an Italian highly specialized hospitals (430 beds). Moreover, during this period, the warehouse and the entry have been involved in the building site. More in detail, the project had a significant decrease in storage

*Corresponding Author Carlotta Patrone, E.O. Ospedali Galliera (Italy), patronecarlotta@gmail.com, +390105632975

space and access gates for goods, medicines and medical devices to the hospital had declined dramatically with an intense impact on the road traffic. In addition, there was no consensus among the management on how to deal with this problem. Furthermore, a great tension between engineers and warehouse managers made it impossible to find a solution. No study in a similar setting has been found in the literature. Therefore, it was necessary to analyze the most appropriate solution for the warehouse management during the construction phase to give to the management all the elements to choose the best solutions.

A deep analysis has been performed and then structured into two different phases, according to Business Process Reengineering (BPR). This paper is organized into five sections. Section 2 provides a literature review of the three main techniques used to develop this study. The methodology followed during the new building construction has been introduced in section 3. Section 4 illustrates the results obtained whereas the conclusions have been presented in section 5.

2. Literature Review

A review of the literature has been performed through the following keywords: "supply chain resiliency", "lean healthcare" and "lean warehouses".

2.1. Supply Chain Resiliency

Supply Chain Resiliency has been defined in a multidisciplinary view as "the adaptive capability of the supply chain to prepare for unexpected events, respond to disruptions, and recover from them by maintaining continuity of operations at the desired level of connectedness and control over structure and function" [3]. From a different point of view, another definition is "The ability of a system to return to its original state or move to a new, more desirable state after being disturbed" [4]. A resilient supply chain must respect the following principles:

- Resilience: it must be part of the initial project and the base of any decision. In other words, it has to be transversal to each side of the company. It is essential a high grade of risk collaboration and the resulting agility in the processes and their connection.

- Reactivity: The company must be ready to quickly respond to the unpredictable events in the most optimal way possible [5]. An efficient company has to work continuously on the vulnerability reduction, decreasing the probability of breakage and enhancing the resilience of their processes. This may occur by increasing the redundancy (increasing the cost) or the flexibility based on the organization of day-to-day operations [6]. Frequently the decision concerning the use of the resilience engineering method and Lean technique is doubtful [7].

2.2. Lean Warehouse Management in healthcare

The Lean has found widespread use in the healthcare sector. Its first hospital application occurred in 2001 in the United Kingdom and the next year in the United States [8]. In 2009, in Canada, five organizations found a way to increase the organization with outstanding results thanks to Lean. Consequently, it has been possible to decrease the waiting time both for the emergency rooms and for medical visits. Moreover, the operating and radiological room utilizations has been enhanced as well as the infectious diseases control [9].

Lean is an appropriate technique to tackle the materials inside the hospitals. A comparison study between the traditional storage and the Lean approach applications has been performed. The collecting data through questionnaires has been used and further statistical analysis has been performed [10]. Another suitable technique is the VIC strategy. This method is based on a mapping of the values flow (V), an integration of the stock information (I) and a final control (C) on these [11]. The flow mapping follows the production path from its start to its end and it shows a visual representation of the material flows and their information in each process [2].

In order to obtain better inventory management performances, the Customer Managed Inventory / Inventory Managed Inventory method (CMI/VMI) could be introduced inside the hospital. VMI/CMI are cooperation forms between a supplier and a client [12].

An improvement system of the warehouse management has been introduced in a hospital in Michigan and it is called "two-bin supply chain system". For each specific medical supply, two containers were provided. This system has been designed in order to reduce the necessary time for the warehousemen to perform the inventory and to decrease the trends of an excessive order [13]. Still inside the hospital warehouse a comparison study between the 5S traditional and the 5S hybrid process has been implemented. 5S is a sequence of activities (sort, set in order, shine, standardize and sustain) and it is demonstrated as an effective tool for the supply management or organization. On the other hand, the hybrid 5S contains inventory management and process improvement tool techniques [1]. In order to manage the warehouses in a quicker and more accurate way it is possible to reinforce the Lean techniques using Radio Frequency Identification (RFID) technologies [14].

Studies focused on the warehouse personnel and on their sensitization about the downtimes have been treated. This waste of useful time could be used for other important procedures inside the hospital [10]. In order to sensitize the personnel, it is possible to introduce the Continuous Positive Reinforcements (CPR) technique. This is a management system using continuous positive reinforcement to achieve the objectives of the working groups [15].

The hospital warehouse management through the Lean application has been successfully described in study developed in province of Parma [16]. Two cases have been analyzed. The purpose of the first one was the cooperation among the hospitals in the region, including an exchange of drugs, medical devices etc., utilizing the same archive software in order to optimize the reorder and the storage process inside each hospital. The second case analyzed was an external centralized warehouse for all the hospitals in the province [16].

Another system adopted in China is based on a partial externalization of the hospital warehouses. It's called the inventory

pooling. In other words, the hospital maintains the inventory minimum level, normally for 1 or 2 days of demand while most of the inventory is retained in the distribution center. The external warehouse organizes the daily deliveries of the drugs based on the information shared real-time by the hospital [17].

2.3. Business Process Reengineering (BPR)

The BPR has spread above all in the Anglo-Saxon world in the early 90s in the private firm sector. In those years, one of the first definition of process reengineering has been provided by Hammer in a paper appeared on Harvard Business Review [18]. In this definition, the author defines this technique as "the substantial afterthought and the radical redesign of the business processes for the purpose to obtain relevant results in terms of quality and services improvements to the user and enhancement in the productivity (cost reduction in view of a volume rise of the performance provided)" [19]. Later, around the middle of 90s, some publications based on this approach started also in the healthcare sector. The BPR in a managerial approach for the management of the organizational change based on the logic for processes.

3. Methodology

One of the key success factors of the BPR implementation is the human involvement [20]. For this reason, a dedicated hospital Lean group has been created. The group is composed of warehouse managers, industrial engineers, civil engineer, architect, nurses, physician, IT technician, and pharmacists.

The authors followed the BPR steps.

3.1. AS IS

The first step was to detect the actual warehouse management with the relative data. The Lean group mapped the flow through the Value Stream Mapping (VSM) technique. Thereafter, a data collection has been implemented in order to detect:

- actual storage space and the number of goods access available;
- real utilization of the dedicated space and of the access.

The actual storage space is 1.631, 25m2 with two dedicated goods access available. It seems to be an easy task but the actual storage spaces are scattered and they are managed by different heads.

For the second purpose a six weeks data detection on storage volume trend and truck arrival has been carried out.

For an accurate data collection, the warehouses have been divided in different areas. Moreover, a monitoring of the truck arrival and departure times, the volume and the timing of the merchandise arrival and the various necessary controls have been performed. Microsoft Excel tables have been created to achieve an efficient collection.

Figure 1 shows a focus on the volume storage trend and Figure 2 illustrates the comparison between standard deviation, mean and the true value of the volume.

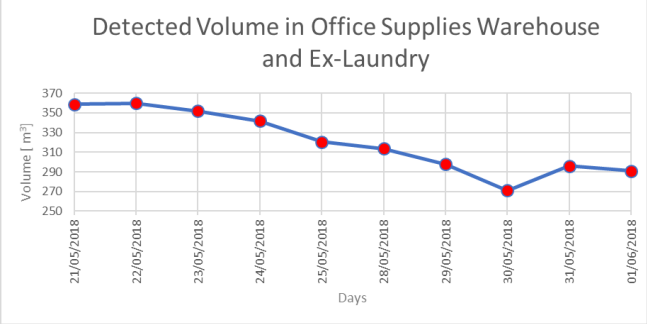


Figure 1: Storage in Office Supplies Warehouse and Ex-Laundry.

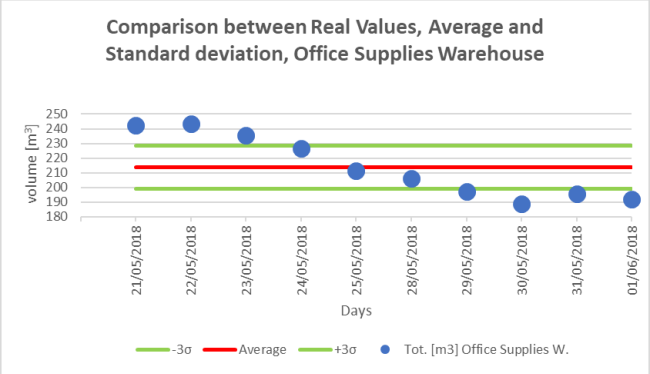


Figure 2: Comparison among Real Values, Average and Standard Deviation, office supplies warehouse.

An additional study based on the filling percentage of the space has been performed and reported in Table 1. This means that many area and shelves do not occupy most of the available space. In parallel to this phase, the template for the arrival of trucks has been filled in. Two different good access are available: one dedicated to the office supply warehouse and other to the pharmacy.

Table 1: Percentage of filling, current state of the warehouse

	[m^3] occupied	[m^3] available	% occupied
Shelves	80,09	273,15	29%
Corridor and Atrium	20,84	74,54	28%
Pallet Rooms 003-007-009	140,41	288,54	49%
Ex-Laundry Room	116,181	398,92	29%

A statistical analysis has been performed to understand if the time slot or the day of the week or the week affect the truck arrival. Latin Square is applied when disturb factors occur in addition to the treatments and to the first disturb factor [21].

Concerning the data analysis of the truck monitoring of the office supply warehouse, the results of the Latin Square Method highlighted a significant difference in the truck arrival by changing the time slots. Day and week did not affect the truck arrival. The results are reported in Table 2.

Table 2: Latin Square Method

Source of Variation	F_0	$F_{0,05;4;12}$	Conclusions
Weeks	0,81	3,26	$F_0 < F_{0,05;4;12}$
Time slots	7,70	3,26	**$F_0 > F_{0,05;4;12}$**
Days	2,27	3,26	$F_0 < F_{0,05;4;12}$

Analysis of Variance (ANOVA) has been performed on that data because only time slots affected the results. ANOVA is a set of statistic techniques whom are part of the inferential statistic. Moreover, they allow to analyze two or more groups of data by comparing the intern variability of these groups with the variability among these [22]. The ANOVA results for the office supply warehouse has been reported, as an example, in the Table 3 and 4. The two most important factors in using ANOVA to accept or reject the null hypothesis were calculated.

Table 3: Anova Results of Office Supply Warehouse (1/2)

Source of Variation	SS	Gf	MS
Time slots	137,62	10	13,762
Days	7,03	4	1,758
Interaction	34,10	40	0,853

There is a statistical significance to the results if F-value (F) is larger than the F critical value (F crit). Also, these results showed that only time slots have the statistical significance as shown in Table 2. The p-value is less than 0.05, so was rejected the null hypothesis that there's no difference between the means and the conclusion is that a significant difference does exist.

Therefore, the One-Way ANOVA test has been performed to understand which time slots are relevant and to provide a ranking of them. The Duncan test has been carried out to define the order of importance among the times using the One-Way ANOVA results. The output of this test, shown in Table 5, has defined which time slots have significant differences to each other.

Table 4: Anova Results of Office Supply Warehouse (2/2)

Source of Variation	F	P-value	F crit	Conclusions
Time slots	**14,92**	3,4E-21	**1,87**	**F>Fcrit p-value< 0,05**
Days	1,91	0,11	2,40	F<Fcrit p-value> 0,05
Interaction	0,92	0,60	1,44	F<Fcrit p-value> 0,05

Table 5: Duncan's Multiple Range Test

Time slots	Treatments (significant differences in bold)	Differences between means	Shortest significant range
8:30-10:00	**6:30-8:30**	8,50	3,19
	10:00-10:30	7,50	3,53
	11:00-12:00	5,33	3,49
	10:30-11:00	3,50	3,41
10:30-11:00	**6:30-8:30**	5,00	3,53
	10:00-10:30	4,00	3,49
	11:00-12:00	1,83	3,41
11:00-12:00	6:30-8:30	3,17	3,49
	10:00-10:30	2,17	3,41
10:00-10:30	6:30-8:30	1	3,41

Different results are achieved by the pharmacy access. Indeed, there are no remarkable differences among the time slots, the days or the weeks there. The truck arrival has a rectangular distribution since the analysis shows that the randomness with which the couriers deliver is very high. Hence, the truck arrival in the drugstore warehouse is a completely random variable with a rectangular distribution.

3.2. TO BE

The future state is the warehouse management with a storage space of 618,8 m^2 and with only one good access for all the hospital. Indeed, during this phase, this gateway will be the unique access to the hospital instead of the actual three (two dedicated to the goods and one two the services).

The authors individuated three possible solutions:

1. *Scenario 1*: internal management of the warehouses.
2. *Scenario 2*: assigning the warehouses management to an external firm of logistic.
3. *Scenario 3*: renting of the new warehouse and the procurement of the logistic of the deliveries to the Hospital.

For all the three solutions the following parameters, reported in section 4, have been studied:

- Economic Analysis: the balance between the emerging and the loss costs.
- Impact on Traffic
- Advantages: the point of strength of the scenario.
- Disadvantages: the point of weakness of the scenario.

Scenario 1

This scenario consists in the completely internal warehouse management. There are two main problems in this context. The first is the 62% less of storage space and the second the unique access to the hospital.

A data extraction on a one-year time frame has been performed in order to calculate the rate of rotation of the goods. The future available space turned out much smaller than the actual one but, thanks to the results performed with the rotation index, it is possible to notice that by increasing this one all the material currently available in the medical devices and office supply warehouse can be incorporated. As shown in the Figure 3, the rotation index of many products is very small or even zero, this means that in a one-year time slot the products are not completely used and their supplies are left. After a Pareto Analysis is possible to establish which products can be moved more often, in order to reduce their stock in the warehouse and to occupy less volume accordingly. It is also possible to determine which products could be put in the higher and less reachable shelves and which ones could be put in the reachable shelves.

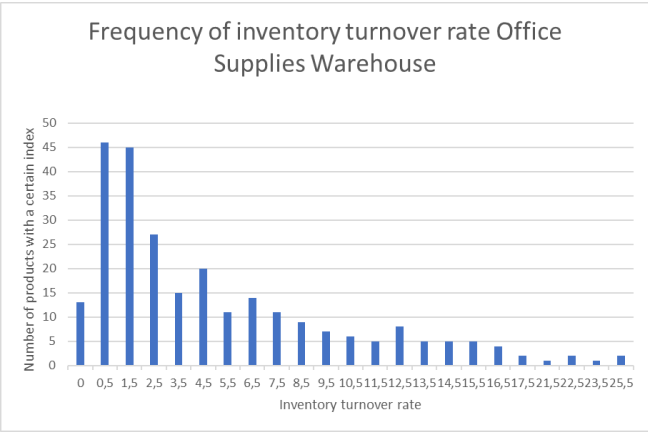


Figure 3: Inventory turnover rate, Office Supplies Warehouse.

Indeed, by increasing the filling percentage in the future shelves, it will be possible to optimize the spaces available and all goods, except drips (IV), can be allocated to warehouses despite the large reduction in space. The IV management will have to be outsourced.

Considering only the contemporary truck arrival of the office supply warehouse and drugstore it is possible to suppose a co-presence of about eight trucks in the rush hours. The future truck check-in ramp during the building of the new hospital will receive all the stocks instead of the actual organization. Furthermore, it will have to be the transit point toward many other stocks directed to other wards (e.g. radiopharmaceuticals, catering services and construction site). In order to understand more about the future traffic, an overlap of the trucks has been carried out and has been reported in Figure 4.

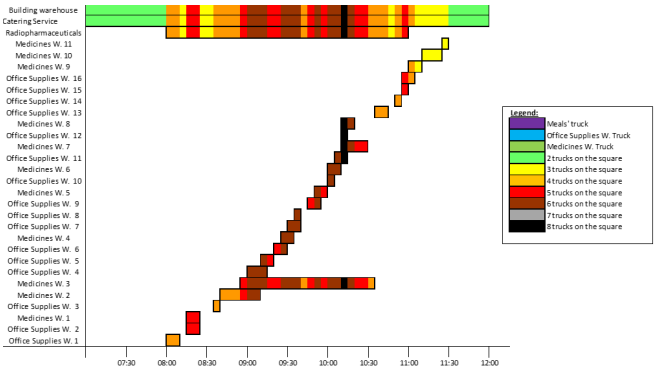


Figure 4: Occupation of the passage square, transitional state.

Figure 4 highlights the co-presence of at least two trucks. In addition, the street on this entrance is a one way one so this is a critical issue to consider.

All the costs involving in this scenario such as the costs for structural costs, outsourcing service for IV, etc. are reported in section 4.

Scenario 2

This scenario consists in the outsourcing of the service. This option comprises the delivery of the goods directly to the wards from the external warehouse. There would not be impact on traffic. Indeed, the goods delivery would happen once a day, at the time chosen by the hospital. The delivery time slot would be probably in the afternoon when the most part of the accesses to the hospital have already been done.

A market analysis has been carried out to understand the cost of this service. In addition, the impact on this scenario on the personnel actually dedicated to the warehouse has been studied with a reduction of them through:

- Pensions
- Termination of fixed-term contracts
- Transfer of personnel

Scenario 3

In this solution the hospital should rent a warehouse and the service of delivery of material from the rented warehouse to the Hospital. A market analysis has been carried out to understand the

cost of this service. The warehouse would be managed the hospital with a staff increase.

4. Results

The detailed results of the cost analysis implemented are reported in Table 6.

Table 6: Scenario Costs Analysis

Cost	Scenario 1	Scenario 2	Scenario 3
Staff	2.700.000€ (5 years x 30.000€/year x 18 workers)	750.000€ (5 years x 30.000€/year x + 5 workers)	2.700.000€ (5 years x 30.000€/year x 18 workers)
Additional or diminutive Staff	7500.000€ (5 years x 30.000€/year x 5 additional workers)	-1.9500.000€ (13 workers less*5years*30.000€/year)	750.000€ (5 years x 30.000€/year x 5 additional workers)
Construction costs	€ 823.858,00	€ 105.000,00	€ 275.150,65
Automated warehouses for drugs			
Wardrobes and shelving warehouse	€ 20.292,65		
I.V. outsourcing	500.000€ (5 years x 100.000€/year)		
Optical pen	€64,66		€64,66
Outsourcing fee		5.000.000€ (5 years x 1.000.000 €/year)	
Rental warehouse			318.000€ (5.300€/months per 60 months)
Rental delivery service from rented warehouse to hospital			2.122.800€ (348.000 €/years*1,22 VAT*5 years)
Total costs (5 years)	**€ 4.794.215,31**	**€ 3.905.000,00**	**€ 6.166.015,31**

Table 7 highlights the advantages and disadvantages of each scenario.

Table 7: Scenario Analysis

SCENARIO	ADVANTAGES	DISADVANTAGES
SCENARIO 1: The exclusively internal management	The directly availability of the stocks. Tendering procedure is not requested.	The rooms' utilization because they are very small. The manual handling is not optimal, the workforce is expected to increase. The medicines' reception and control will be problematic because of the questions that arise about the responsibility and the increase of staff.
SCENARIO 2: The exclusively external management	The solution is immediate	Tendering procedures and the consequently long time
SCENARIO 3: the partial use of internal structures and the rent of the external warehouse	The direct control over the material. There is no immediate expense for the renovation of the space necessary for the transitional warehouse.	Tendering procedures. The increase of the complexity in the management of remote staff and the probable increase of employees.

5. Conclusions

The completely intern solution (Scenario 1) has a very strong impact on the viability. One advantage of this solution is that the items are immediately available and there are no conflicts in the item management, shorter availability time and no tender procedure. The uniquely external solution (Scenario 2) has a low impact on the viability and it is immediate. There is a loss of control in the stocks, confusion in the role of the stock responsibilities etc.

The partially intern solution (Scenario 3) with the rent of an external warehouse has a low impact on the viability. The advantages are the direct control on the merchandise, lack of an immediate disbursement for the restructuring of the spaces. On the other hand, the drawbacks are the increasing in complexity of the remote employee management and a probable increase in personnel.

All the analysis has been subjected to the attention of the Direction. They decided to adopt Scenario 2.

This paper reports a method and it could be applied in all the contexts during the building of a new construction. In the hospital this approach has been used because the management didn't know how to manage the stocks during the building of the new hospital

because the spaces for the warehouse were limited and there was an impact on the traffic.

The authors implemented the BPR approach involving all the professionals. This approach achieved the goal because it gives all the elements to the Management to make a decision and all the people feel involved in this phase. In fact, at the end of the project the tensions initially present were completely eliminated.

One possible limitation of the study could be that the market analyses carried out on the August of 2018 could be may change depending on new solutions provided by companies. However, the management of the hospital needed to take a decision on that moment. Nevertheless, the aim of this paper is to provide a method and not an economic estimate. Future research on this topic could be done using discrete event simulation to have more detail on the impact of each solution on the system.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

A special thanks to Maria Laura Zizzo, Cinzia Gina Cravisana, Ezio Nicolas Bruno Urbina. Federica Solari, Robrta Clavarino, Mauro Nelli and Carla Elda Fraguglia for their important contribute to this paper.

References

[1] S. Venkateswaran, "Thesis. Implementing lean in healthcare warehouse operations - evaluation of 5S best practice" Louisiana State University and Agricultural and Mechanical College Master's Theses, 2011.
[2] P. Frank, C. Garcia, "Applying Lean Concepts in a Warehouse Operation" IIE Annual Conference and Exhibition. 2004. p. 2819-2859.
[3] S. Holcomb, "Understanding the concept of supply chain resilience" The international journal of logistics management, **20**(1), 124-143, 2009.
[4] Collins, H. (2000). Collins English Dictionary. Glasgow.
[5] H. Peck, M.C. Peck, "Building the resilient supply chain" The international journal of logistics management, **15**(2), 1-14, 2004.
[6] Y. Sheffi, J.B. Rice Jr. "A Supply Chain View of the Resilient Enterprise" MIT sloan management review, **47**(1), 41-48, 2005.
[7] T. A. Saurin, C. B. Rosso and A.L. Colligan "Towards a resilient and lean health care," In: Resilient health care. ROUTLEDGE in association with GSE Research, 2017. p. 3-17.
[8] H. Yalin, and T.E. Yonetimi, "Lean Medical Inventory Management in Hospitals," International Journal of Academic Value Studies (Javstudies), **3**, 16-24, 2017.
[9] B.A. Fine, B. Golden, R. Hannam and D. Morra, "Leading Lean: A Canadian Healthcare Leader's Guide," Healthcare Quarterly, **12**(3), 32-41, 2009.
[10] A. Yildirim, "Innovative approach to health service materials management" Hestourex World Health Sports Tourism Congress & Exhibition, 6-9 April 2017, Antalya (33-42).
[11] E.F. Ramly, "Lean Supply Chain contribution toward Lean Health Care," International Conference on Continuous Value Adding based on Lean Thinking in Hospital Management, Tehran, Iran, 30 Nov – 1 Dec 2011.
[12] L. Michailidou, "Healthcare supply chains: a case study of hospital vendor collaborative practices" International Journal of Logistics Systems and Management, **15**(2), 288-303, 2013.
[13] E. Aljure, L. Busse, J. Hahn and S. Hayes, "Standardizing and improving the efficiency of Michigan Medicine's supply chain management of hospital supplies," 2017. http://umich.edu/~ioe481/ioe481_past_reports/17W01.pdf
[14] J.C. Chen, C.H. Cheng, P.B. Huang, K.J. Wang, C.J. Huang and T.C. Ting, "Warehouse management with lean and RFID application: a case study," he International Journal of Advanced Manufacturing Technology, **69**(1-4), 531-542, 2013.
[15] J. Benefield, "CPR For Lean Warehouses," Madison: Advantage Business Media. M., 2015. https://www.inddist.com/operations/article/13763532/cpr-for-lean-warehouses
[16] M. Bertolini, F. Bevilacqua, E. Ciarapica, and L. Postacchini "Business process reengineering of drugs storage and distribution: a case study" International Journal of Procurement Management, **8**(1/2), 44-65, 2015.
[17] D. Wua, J.E. Tepper, "Possibility of Inventory Pooling in China's public hospital and appraisal about its performance" Applied Mathematical Modelling, **39**(23-24), 7277-7290, 2015.
[18] M. Hammer, "Reengineering work: don't automate, obliterate," Harvard Business Review, **68**(4), 104-112, 1990.
[19] M. Motta, F. Lega, "Process re-engineering per le aziende sanitarie: una introduzione," Mecosan, **6** (23), 163-165, 1997.
[20] M. Al-Mashari, M. Zairi, "BRP implementation process: an analysis of key success and failure factors," Business process management journal **5**(1), 87-112, 1999.
[21] P.R. Mosca, "Il progetto quadrato latino," University of Genoa - Dipartimento di Ingegneria della Produzione - Gestione degli Impianti Industriali I
[22] D. Montgomery, Design and Analysis of Experiments (5th ed.). New York: Wiley, 2001, ISBN 0-7167-9657-0.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 263-266 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Design and Optimization of ZnO Nanostructured SAW-Based Ethylene Gas Sensor with Modified Electrode Orientation

Argha Sarkar[1,*], Padarthi Venkataramana[1], Nimmala Harathi[2], Thummuru Jyothsna[1], Neeruganti Vikram Teja[1]

[1]*Nanoelectronics Lab, Department of ECE, Sree Vidyanikethan Engineering College, 517102, India*

[2]*Nanoelectronics Lab, Sree Vidyanikethan Engineering College, 517102, India*

ARTICLE INFO

Article history:
Received: 01 November, 2019
Accepted: 19 January, 2019
Online: 27 January, 2020

Keywords:
Surface Acoustic Wave,
Hexagonal ZnO Nanorod,
Interdigitated Transducer,
Sensitivity.

ABSTRACT

Here the work approaches the detection of fruit maturity level. In addition, a gas sensor is proposed to obtain better sensitivity to know the level of ethylene gas released from matured fruit. Better sensitivity can be achieved by obtaining an enhanced active area, changing the intermediate layer thickness and the pitch value of the electrode. Hexagonal nanostructured zinc oxide is used as a sensing layer. Finally, frequency shift and a total displacement of the device are increased. The hexagonal structured ZnO nanorod-based surface acoustic wave gas sensor may be used to detect ethylene gas in fruit ripening sensor with enhanced sensitivity. This paper reports a simulation of Surface Acoustic Wave (SAW) wave propagation with respect to interdigitated transducer (IDT), piezoelectric material and sensing layers.

1. Introduction

It is becoming very important to design a highly sensitive fruit monitoring system to determine the storage life of fruit and quality. Based on the maturity level of fruit the decision may be taken regarding transportation, consumption, etc. The ripe fruits are very tough to handle as they become soft and squashy with an insipid smell and they cannot withstand transportation from one place to another. Hence it is important to monitor the state when they are mature. Skin color, hardness, size, and smell are the common parameters to get an idea about the fruit ripening. But these are basically assumption-based assessment, so it is havening an issue of reliability [1].

Ethylene (C_2H_4) is one of the organic gases which is formed in the ripening process of climacteric fruits. These fruits (eg. Mango, guava, banana, apricot, kiwi, palm, etc.) liberate C_2H_4 during ripening process [2-4]. Climacteric fruits pass through climacteric state after harvesting. A small dose of C_2H_4 in a controlled environment can simulate the ripening process of those climacteric fruits. So, if the maturity levels of these fruits are detected efficiently then they can be easily stored or transported after harvesting. On the other hand, non-climacteric fruits [5] emits a little amount of C_2H_4 and they are in a position to harvest only when completely ripened.

Herein, the work focuses on the development of surface acoustic wave-based ethylene gas sensor for detection of fruit ripening at different maturity stages. Detecting the C2H4 gas concentrations that has been liberated by fruits, may evaluate the maturity stage and the shelf span can be estimated [6]. In this study, a SAW-based sensor is designed and sensitivity is analyzed. Interigitated transducer (IDT), piezoelectric material and sensing layers are a major consideration while designing a SAW sensor. Pitch is one of the important parameters to determine the synchronous frequency. On the surface of the piezoelectric material, aluminum IDT generates acoustic wave from electrical signal by producing a mechanical force [7-8]. Adsorption of target gas molecules take place on the sensing material, which happens between the two IDTs. ZnO nanorod based SAW sensor is designed and wave properties are studied. Parameters like device length, height, intermediate layer thickness, and IDT thickness are optimized through simulation study to obtain enhanced sensitivity.

2. Sensor's Working Principle

The template SAW gas sensors are dominant from other sensors because of detection of physical, biological, and chemical quantities. It is only due to minimized insertion loss, increased sensitivity, and gas detection capability. Their fundamental operation is based on the changes in a propagating acoustic wave due to the target gas interaction, which depends on the the

*Corresponding Author: Argha Sarkar Email: argha15@gmail.com

propagation medium, mass density, electric–dielectric behavior and the elastic stiffness of the piezoelectric materials. Due to the piezoelectric effect, electrical signal is generated when the acoustic wave travels from the surface of the device substrate to an IDT. In SAW based gas sensor, a piezoelectric material plays a major role as a transducer. It transforms a signal from electrical to mechanical wave by utilizing the piezoelectric effect. Piezoelectric effect relates structural mechanics with the electrostatic field. The electric charge is created when a sinusoidal signal is applied to a piezoelectric material. The properties of a piezoelectric material rely on some important parameter like electric flux (D), resonating frequency (f_0), and stress (T). The functionalities of the SAW device are highly dependable on those parameters. To obtain the electrostatic fields (D) [9] from the equation 1, it is important to know about the parameters like stress vector (T), strain vector (S), electric field (E), dielectric matrix (e), piezoelectric matrix (ε) and elasticity matrix (c).

$$D = e^T S + \varepsilon E \tag{1}$$

From equation 1 it is cleared that it is required to know the structural mechanics of piezoelectric material [7,10]. Equation 2 is to find out the value of T.

$$T = cS - eE \tag{2}$$

Device efficiency is high when input signal frequency matches with resonating frequency. The resonating frequency of a SAW device is associated with propagation velocity (V_p) of the acoustic wave and the pitch value (p) (shown in equation 3 [11-12]).

$$f_0 = V_P / p \tag{3}$$

As per the equation 4, the central wavelength SAW sensor can easily be found from the electrode pitch (p) and finger width (h).

$$\lambda = 2(h + p) \tag{4}$$

IDT plays a vital role to obtain enhanced SAW sensor response. The lengths of the IDT, finger width, pitch are the important considerations. Bandwidth (B), acoustic wave frequencies are decided based on input IDT. For a specific operating frequency, the range of acoustic wave frequencies, can be enhanced by optimizing the number of pair of fingers (N) in IDT. Mathematically the bandwidth can be defined as [7,11],

$$B = \frac{2 f_0}{N} \tag{5}$$

If the number of pairs of finger in optimized the length of IDT (l_{IDT}) will also get changed.

$$B = \frac{2V_P}{NP} \tag{6}$$

$$B = 2\frac{2V_P}{l_{IDT}} \tag{7}$$

A sharp signal may be obtained by minimizing the bandwidth [11]. From equation 7 it can be said that bandwidth is inversely related to the length of the IDT. The sensing mechanism is mainly based on the adsorption phenomenon. When the sensor is placed in the target gas environment, target gas molecules are adsorbed on the surface of the sensing layer. It results in a change in the wave velocity. Because of the mass loading effect [13-14] there must be a frequency shift (given in equation 8) by which the detection can be performed.

$$\Delta f = (K_1 + K_2) f_0^2 \sigma - K_2 f_0^2 \sigma \frac{4\mu}{V_0^2}\left(\frac{\lambda + \mu}{\lambda + 2\mu}\right) \tag{8}$$

$$\Delta f = (K_1 + K_2) f_0^2 \sigma \tag{9}$$

(By neglecting the right part of equation 8)

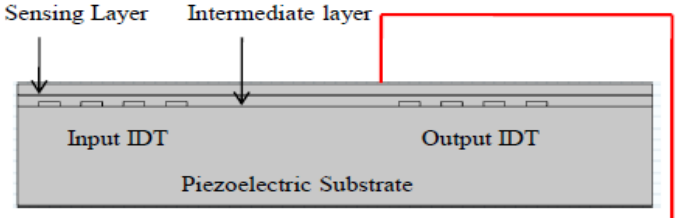
Where σ is the areal density of the layer after gas adsorption and λ, μ are signify Lame constant and shear modulus of the film respectively. If the value of Δf is low, there will be no significant change in frequency. It is not desired for a highly sensitive sensor. Δf is required to be significantly higher value. From equation 8 it can be said that $[K_2 f_0^2 \sigma \frac{4\mu}{V_0^2}(\frac{\lambda + \mu}{\lambda + 2\mu})]_{Min}$ gives the maximum frequency deviation. For a piezoelectric material $LiNbO_3$ and sensing layer ZnO, the value of shear modulus related expression becomes so small that $[K_2 f_0^2 \sigma \frac{4\mu}{V_0^2}(\frac{\lambda + \mu}{\lambda + 2\mu})]$ can be ignored [14].

The reduced expression is shown in equation 9.

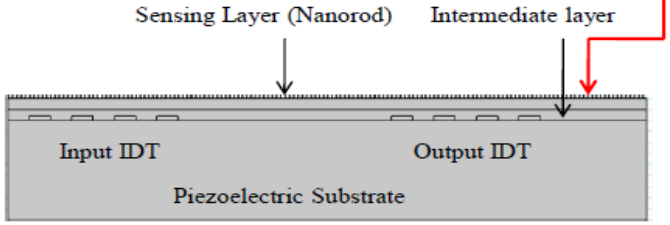
current designations.

3. Sensor Design and Simulations

Comsol Multiphysics 5.5 ® is considered to design and simulate the SAW-based gas sensor. Herein, initially simple ZnO based SAW gas sensor is considered and then it is embedded with ZnO nanorods of the surface of the ZnO sensing layer. Optimization of the delay-line two port configuration and applied frequency results in device performance. The important parameter of piezoelectric material lithium niobate is having a higher value of coupling coefficient and surface acoustic wave velocity. Lithium niobate is considered as a substrate material to optimize the device. The mass loading effect results after embedding MEMS module on the piezoelectric device. The frequency potential and displacement of the surface can be determined by using COMSOL Multiphysics. Sensing layer plays a vital part to get higher sensitivity. ZnO [15] is found to be an appropriate material towards gas sensing and it is also available in the list of built-in material in this simulation environment.



(a)



(b)

Figure. 1 Side view of the layout of (a) layered (b) Hexagonal nanostructured SAW gas sensor

The dimensions of the IDTs can be calculated by using equation 4 of 100 MHz signal frequency. Figure 1 a. shows the layout of the layered SAW gas sensor. Hexagonal nanorods are added on the nanolayered structure to obtain enhanced sensitivity. The sensitivity study is done based on the effect of hexagonal nanorods. In both cases, the simulation study is done with different values of device length, height, intermediate layer thickness, and IDT thickness. Here Figure. 1 (a) and (b) are the optimized design layouts.

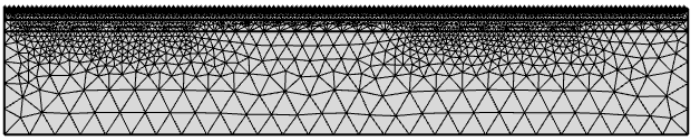


Figure. 2 2D meshing analysis of a SAW Sensor as the sensing material with nanorods

Where device length is 30 µm and its height is 4 µm. The thickness of the intermediate layer is 0.2 µm and ZnO layer is of 0.5 µm. The aluminum is taken as IDT material. The width of the IDTs is 1µm and height 0.1µm. To avoid the second order effect, electrodes are considered as a massless electrode. Figure. 2 represents the two dimensional cross sectional layout after doing the meshing. The node size is taken in such a way so that it becomes less than the wavelength and compensates the appropriate harmonics. Figure. 2 also ensures that the maximum number of nodes present is below the IDT location.

Where device length is 30 µm and its height is 4 µm. The thickness of the intermediate layer is 0.2 µm and ZnO layer is of 0.5 µm. The aluminum is taken as IDT material. The width of the IDTs is 1µm and height 0.1µm. To avoid the second order effect, electrodes are considered as a massless electrode. Figure. 2 represents the two dimensional cross sectional layout after doing the meshing. The node size is taken in such a way so that it becomes less than the wavelength and compensates the appropriate harmonics. Figure. 2 also ensures that the maximum number of nodes present is below the IDT location.

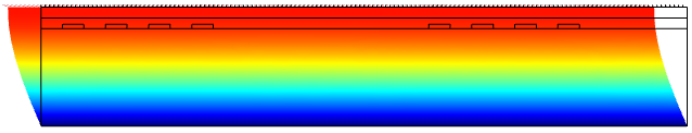


Figure. 3 Simulation of SAW Sensor as the Sensing Material with nanorods

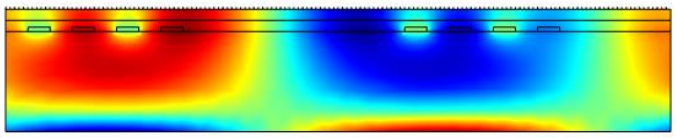


Figure. 4 Simulation of SAW Sensor as the Sensing Material with nanorods after exposing to gas

During the operation of the SAW device, a 150 Hz sinusoidal signal is given to the input IDT. It results in an electric field between the two opposite polarity IDT fingers. As the substrate is the piezoelectric material, due to the piezoelectric effect tensile and comprehensive strain will be created. Thus mechanical wave will appear. The wave will propagate through the surface of the material. Reflector and absorber are used to minimize the effect of insertion loss. When the analytes (ethylene) comes in contact with the sensing platform, sensing layer adsorbs the gas and results in increment in resonant frequency. Here, the sensing layer acts a vital role. By achieving higher surface to volume ratio of the nanostructured sensing material it can be expected that the adsorption of the target gas may be increased which in turn results in enhanced sensitivity. In this simulation study, the presence of ethylene gas is detected through the displacement in resonant frequency. Once the ethylene is detected with higher sensitivity, it will become easy to ensure the ripening state of the fruit.

4. Result and Analysis

After meshing and simulation (shown in Figure. 2 to Figure. 4) different colors have appeared. It is due to the applied voltage. More energy is confined above at the surface of the piezoelectric material [16]. Surface acoustic wave can be seen at the surface of the substrate yielding a noticeably change in propagation properties. The thickness of the sensing layer may effect on the sensitivity. A complete analogy of total displacement with respect to the device is shown in Figure. 5 and Figure. 6 upon exposure to ethylene gas.

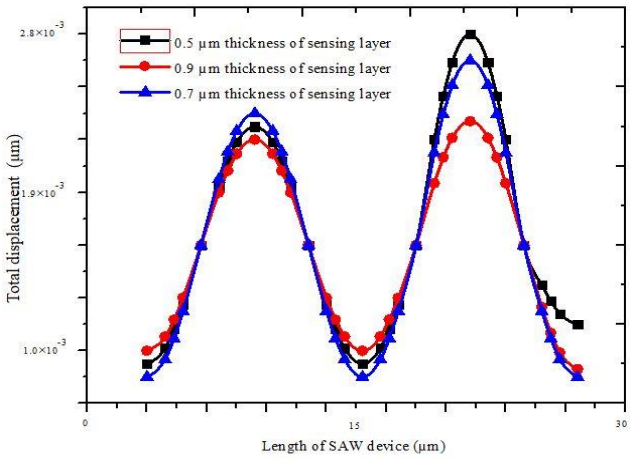


Figure. 5 Total displacements versus device length by varying the thickness of sensing layer

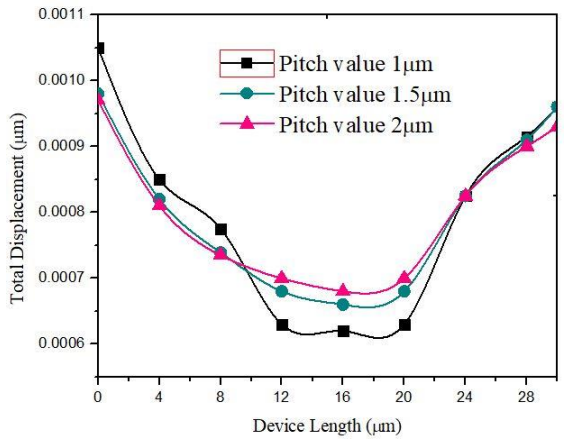


Figure. 6 Total displacements versus device length by varying the pitch value of IDT

Figure. 5 confirms that 0.5 μm thickness of the sensing layer provides increased displacement thus sensitivity. The dependency of sensitivity on pitch value is studied in Figure. 6. Aluminum is chosen as an IDT material because of its low sensitivity and low cost. Different pitch value of 1 μm, 1.5 μm and 2 μm are considered and it is found that 2 μm provides higher displacement of SAW at both input and output and output for a device length ranging from 8 to 20 microns. It results in enhanced sensitivity [17].

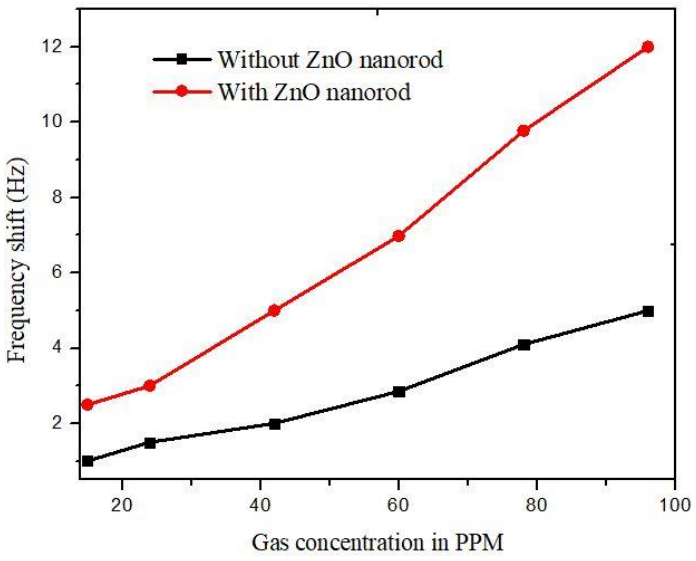


Figure. 7 Frequency shift due to gas exposure in presence and absence of ZnO nanorod

The frequency shift of the sensor is observed because of the different concentration of the target gas. The shift of frequency is present due to C_2H_4 gas release from fruit. The effects of ZnO nanorods on sensitivity are studied in Figure. 7. The presence of ZnO nanorod over the ZnO sensing layer exhibits better frequency shift compared to bared ZnO layered SAW sensor. Simulation results assure that ZnO nanorod based SAW sensor provides higher sensitivity. It is due to increased surface to volume ratio. Hexagonal nanorod provides increased surface to volume ratio i.e. more active region on the sensing layer which results more reaction with target analyte and increased displacement.

5. Conclusion

All An ethylene gas sensor to measure the fruit maturity is designed and simulated in COMSOL Multiphysics to study the results in resonant frequency by using ZnO sensing layer with proper electrode orientation. Furthermore, the ZnO nanorods are considered to obtain better sensitivity. A comparative study based ZnO layered and nanorod based sensors are performed in COMSOL Multiphysics. It reveals that nanorod based sensor provides enhanced frequency shift due to its higher surface to volume ratio in hexagonal structure. The displacement in resonant frequency is because of adsorption of the analyte by the sensing layer. Here future work includes the fabrication of ZnO nanorod and SAW sensor to verify the simulation results.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to acknowledge Sree Vidyanikethan Engineering College for infrastructural and laboratory support.

References

[1] T. Barboni, , M. Cannac, N. Chiaramonti, “Effect of cold storage and ozone treatment on physicochemical parameters, soluble sugars and organic acids in Actinidia deliciosa” Food chem., 121(4), 946-951, 2010. https://doi.org/10.1016/j.foodchem.2010.01.024

[2] S.P. Burg, E. A. Burg, “Relationship between ethylene production and ripening in bananas” Bot. Gaz., 126(3), 200-204, 1965. https://doi.org/10.1086/336320

[3] P. P. Iannetta, L. J. Laarhoven, Medina-Escobar, E. K. James, M.T. McManus, H.V. Davies, F.J. Harren, “Ethylene and carbon dioxide production by developing strawberries show a correlative pattern that is indicative of ripening climacteric fruit” Physiol. Plant., 127(2), 247-259, 2006. https://doi.org/10.1111/j.1399-3054.2006.00656.x

[4] C.S Barry, J. J. Giovannoni, “Ethylene and fruit ripening” J. Plant Growth Regul., 26(2), 143, 2007. https://doi.org/10.1007/s00344-007-9002-y

[5] S. O. Nelson, W. Guo, S. Trabelsi, Investigation of dielectric sensing for fruit quality determination. IEEE Sensors Applications Symposium, February, Atlanta, GA, USA. pp. 54-56, (2008). https://doi.org/10.1007/s00344-007-9002-y

[6] S. Janssen, K. Schmitt, , M. Blanke, M. L. Bauersfeld, J. Wöllenstein, W. Lang, “Ethylene detection in fruit supply chains” Phil. Trans. R. Soc. A: Mathematical, Physical and Engineering Sciences, 372(2017), 20130311, 2014. https://doi.org/10.1098/rsta.2013.0311

[7] J. Kirschner, “Surface Acoustic Wave Sensors (SAWS)” Micromechanical Syst., 1-11, 2010. https://doi.org/10.18689/mjnn-1000109

[8] M. Anand, ed., Study of tin oxide for hydrogen gas sensor applications. Thesis Scholar Commons University of South Florida, 2005.

[9] E. Comini, G. Faglia, G. Giorgio, Solid State Gas Sensing, Springer, 2009.

[10] Ballantine Jr, D.S., White, R.M., Martin, S.J., Ricco, A.J., Zellers, E.T., Frye, G.C. and Wohltjen, H., eds. Acoustic wave sensors: theory, design and physico-chemical applications. Elsevier, 1996.

[11] S. Capone, A. Forleo, L. Francioso, R. Rella, P. Siciliano, J. Spadavecchia, D. S. Presicce, A. M. Taurino, “Solid state gas sensors: state of the art and future activities” J Optoelectron Adv M, 5(5), 1335-1348, 2003. 10.1002/chin.200429283

[12] F. Nawaz, V. Jeoti, “SAW sensor read range limitations and perspectives” Wireless networks, 20(8), 2581-2587, 2014. https://doi.org/10.1007/s11276-014-0765-3

[13] M. N. Hasan, D. Acharjee , D. Kumar, S. Maity, "Simulation of low power heater for gas sensing application" Procedia Comput. Sci., 92 (2016), 213 – 221, 2016. 10.1016/j.procs.2016.07.348

[14] M. N. Hasan, S. Maity, A. Sarkar, C. T. Bhunia, D. Acharjee, A. M. Joseph, "Simulation and fabrication of SAW-based gas sensor with modified surface state of active layer and electrode orientation for enhanced H2 gas sensing" JEM, 46(2), 679-686, 2017. https://doi.org/10.1007/s11664-016-5128-7

[15] A. Sarkar, S. Maity, C.T. Bhunia, P.P. Sahu, "Responsivity optimization of methane gas sensor through the modification of hexagonal nanorod and reduction of defect states" Superlattices and Microstruct., 102, 459-469, 2017. https://doi.org/10.1016/j.spmi.2017.01.006

[16] G. Di Francia, B. V. Alfano, La Ferrara, "Conductometric gas nanosensors" J. Sens., 2009, pp 1-18, 2009. https://doi.org/10.1155/2009/659275

[17] G. Zhang, "Nanostructure-enhanced surface acoustic waves biosensor and its computational modeling" Journal of Sensors, 2009, 1-11, 2009. https://doi.org/10.1155/2009/215085

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 263-269 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Prediction of Demersal Fishing Ground Associated with Coral Reefs in the Coastal Jepara Regency, Central Java, Indonesia Based on Sentinel 2a Imagery

Kunarso[*,1,2], Muhammad Zainuri[1,2], Denny Nugroho Sugianto [1,2], Jarot Marwoto[1], Hariyadi[1], Muslim[1]

[1]*Department of Oceanography, Faculty of Fisheries and Marine Science, Diponegoro University, Semarang, Indonesia, Prof. Soedarto, SH Street, Tembalang Phn./Fax. (021) 7474698 Semarang 50275*

[2]*Center for Coastal Rehabilitation and Disaster Mitigation Studies (CoREM), Diponegoro University, Center of Excellence Science and Technology (PUI), Indonesia*

ARTICLE INFO

Article history:
Received: 11 September, 2019
Accepted: 22 November, 2019
Online: 05 December, 2019

Keywords:
Fishing Ground
Demersal
Sentinel 2a
Jepara

ABSTRACT

Map of prediction of fishing ground that already exists in Indonesia issued by the Bali Marine Research and Observation Center (BPOL) and the National Aeronautics and Space Agency (LAPAN) of Jakarta, there are still many weaknesses including the spatial aspect of the point of forecasting far from the coast and the unclear type of fish predicted. The lack of understanding of the type of fish predicted, the risk of fishermen carrying the wrong fishing gear so that when reaching the forecasting location, fishermen can not catch fish that are abundant. The method of determining fishing ground for demersal fisheries is carried out in three stages, first step, environmental tabulation of demersal fisheries, the second stage of mapping corals area and its supporting parameters, the third stage of mapping the demersal fishing ground. The results of the study obtained several statements, namely demersal fishing ground can be made based on an understanding of childhood fish habitat, adulthood, substrate base habitat, an optimum depth of catchment area. Sentinel 2A satellite imagery can be used to help determine the fishing ground of demersal fish species, but for the coastal waters of Jepara Regency, there are still obstacles related to spatial resolution and water turbidity. Field data is still needed to help determine the demersal fishing ground. Based on the fishing ground mapping, there are four economical types of fishery resources that live in the coral reef environment, namely Lobster, Red Snappers, Blue Line Seabass and Leopard Coral Grouper. The demersal fishing ground associated with corals in the Coastal Jepara Regency Waters which is seven segments apparently only four segments have the potential to become demersal fishing ground associated with coral reefs, namely the second to fourth segments, including the coastal waters of Jepara, Mlonggo, Bangsri, and Kembang. District and the most potential is the coastal waters of the Mlonggo District.

1. Introduction

Catching systems in Coastal Jepara District are generally still traditional because without the help of high technology, only based on experience or feeling alone. Although the fishing gear used is quite advanced, but the departure to the sea has a very high gambling rate and tends to remain like the catchment area of the previous year. The increasing number of boats, making a fishing ground congested with fishing effort so that the results catch per boat tends to fall from year to year [1]. Fishermen's lack of understanding the temporally and spatially changing of fishing ground due to changes in oceanographic conditions makes fishing more difficult, time and costly-consuming, but the results are less than optimal.

In Indonesia, there are already institutions that make fishing ground predictions, namely BPOL (Research and Marine Observation Center) in Perancak Bali and LAPAN (National Aeronautics and Space Institute), but their utilization of fishing ground prediction is not optimal, because there are still many

*Kunarso, et al., Oceanography Department, Diponegoro University, +6285875762162 . kunarsojpr@yahoo.com

weaknesses that need to be fixed. These weaknesses include locations plotted as potential catchment areas generally in offshore areas so that small fishermen cannot take advantage, forecasting in the form of spots so they do not represent the area, besides that there is no specification of fish species, whether for demersal or pelagic fish species. Different types of fish require different habitat locations and different fishing technologies. Another weakness is the basic data used for forecasting is especially based on two parameters, namely chlorophyll-a (fertility indicator) and sea surface temperature from MODIS imagery, this is certainly not suitable if for forecasting demersal fish [1]

Demersal fisheries types from fishes, crustaceans or mollusks are generally associated with vital habitats in the form of corals, seagrasses and mangroves [2]. Types of demersal fisheries that are associated with coral on the coast of Jepara, for example, lobsters, red snapper, coral grouper and tiger grouper [3]. A good understanding of the location of coral reefs can help identify the demersal fishing ground associated with coral. Sentinel 2 satellite imagery has proven to be very satisfying to be used to identify and monitor coral reefs in coastal waters [4,5] The purpose of this research is to make an innovation in determining the fishing ground for demersal fisheries associated with coral reefs on the coast based on data from sentinel 2 satellite imagery, corals field data, and bathymetry data. Research Location in the coastal waters of Jepara Regency, Central Java Indonesia (Figure 1).

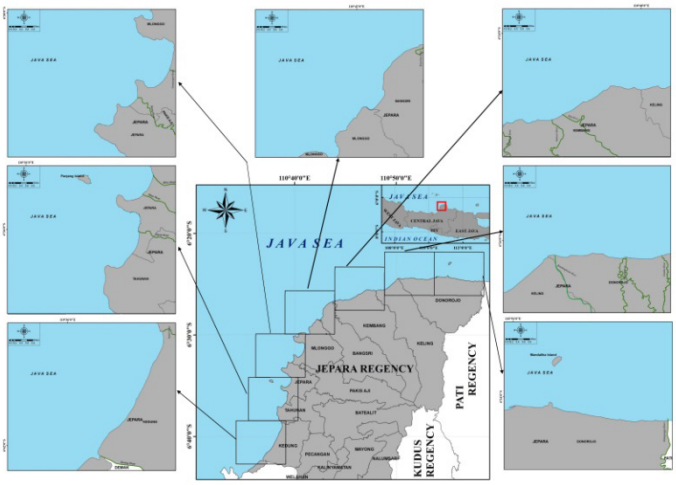


Figure 1: Research Location in the Coastal of Jepara Regency Waters

2. Materials and Methods

2.1. Research Material

The material used in this study includes main data and supporting data. The main data used consists of vital habitat distribution data, namely the distribution of coral reefs, estuarine locations, besides oceanographic parameter data, including bathymetry, total suspended solids (TSS), and bottom substrate. Supporting data include Jepara Regency marine environment map data and demersal fish bioecology data. The data of corals and TSS (Total Suspended Solid) were obtained from sentinel-2 satellite image data processing. The coral's distribution data supplemented with field survey data. Bathymetry data from the Navy Hydro-oceanography Center (PUSHIDROSAL).

2.2. Research Methods

The method used in this study is a descriptive method that seeks to describe the coral's distribution and oceanographic conditions especially TSS that support the life of demersal fish. Furthermore, describe the optimum bioecological factors for demersal fish. The corals, TSS, and the estuary locations are used for delineation (determination) of demersal fish potential prediction in the waters of Jepara Regency.

- *Method of Collecting Data*

The Distribution data of coral reefs from Sentinel-2 Imagery, besides corals data, TSS data were also obtained from Sentinel 2 this has been done by researchers [6,7], corals data from field also be surveyed by acoustic check, and estuarine location information used in this study. The Sentinel satellite data was obtained from the web site www.earthexplorer.usgs.gov. The spacial resolution of sentinel image data of 10 m. The water depth data obtaining from earthexplorer.usgs.gov

- *Data analysis and Processing*

a. Sentinel-2 Satellite Image Data

The Sentinel-2 image processing is carried out in several stages, including image correction and processing and using the lyzenga algorithm to determine the extent of corals. The image used in this study is sentinel-2 which has a cloud cover of less than 10% so that the sea and land look well. Sentinel-2 imagery has 12 bands, in which there are green, blue and NIR channels with various spatial resolutions (10m, 20m, 60m). The required channel in the lyzenga algorithm has been fulfilled.

The first stage of sentinel image processing is geometry correction, this correction is done in the global mapper software and aims to bring up coordinates in the image. The second step is radiometry correction, this correction is done in software mapper and aims to eliminate disturbances when recording images. The third stage is determining the extent of corals using the lyzenga algorithm.

The lyzenga algorithm used [8, 5]

$$Y = \ln(L_i) - [\left(\frac{ki}{kj}\right) x\ ln(L_j)] \quad (1)$$

Where:

L_i = blue channel reflectant value

L_j = the reflectance value of the green channel

k_i/k_j = ratio of attenuation coefficients of blue and green channels

Equation (1) is used for the extraction of water bottom information (Y). That requires two image spectral channels and ki / kj ratio which is the attenuation coefficient ratio between the two channels, in this case, the blue channel and the green channel. This algorithm calculation is influenced by channel pairs i (blue channel) and j (green channel) used. The wavelength of the channel used will affect how deep the channel can detect the

bottom of the water. Blue canals and green canals have the best penetration wavelengths among other channels [5].

The attenuation coefficient (ki / kj) is calculated by the equation:

$$\frac{k_i}{k_j} = a + \sqrt{a^2 + 1} \quad (2)$$

The value of a is determined by the equation:

$$a = \frac{\sigma_{ii} + \sigma_{jj}}{2\sigma_{ij}} \quad (3)$$

Where: σ_{ii} = channel variety or variant i

σ_{jj} = channel variety or variant

b. Sentinel-2 Satellite Diagram of Method for Determination of Demersal Fishing Ground

The innovations to create a potential area for demersal fishing ground, carried out with the stages described in three flow diagrams.

- The first stage

The first stage is the environmental tabulation of demersal fisheries. This stage is carried out with reference studies and collecting catch data from the Jepara Regency area which will be predicted for its fishing ground area. This stage is carried out with a reference study from the Department of Maritime Affairs and Fisheries of Jepara Regency and a survey to the Fish Auction Site (TPI) in Jobokuto Jepara, about the types and amount of demersal fish catches. Then conducted to reference study about the bioecology of demersal fish available from various books and websites. The results of this reference study and survey are tabulated in the environmental table for demersal fisheries.

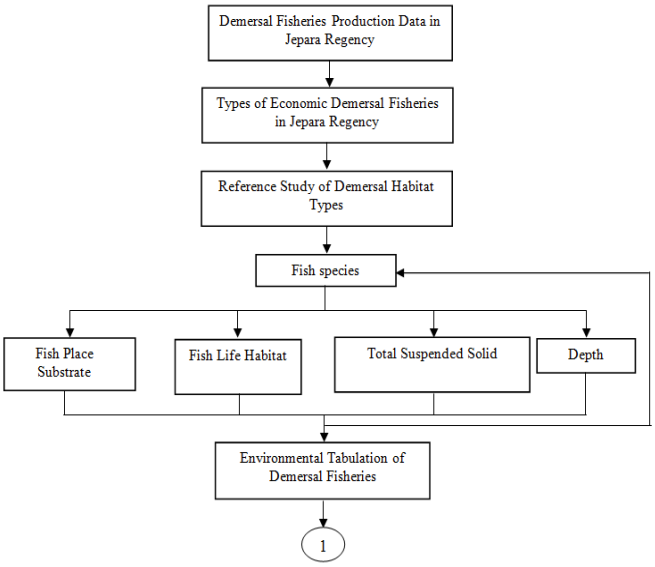


Figure 2: Flow chart of demersal fish environmental tabulation in Jepara Regency

- The second stage

The second stage is the mapping of vital habitat (coral reefs) and its supporting parameters. The second phase is done by downloading Sentinel-2 and bathymetry data, be supported by soft file data of the Jepara Regency environmental map. All data is processed to display the distribution of corals, water depth, TSS and location of river mouths.

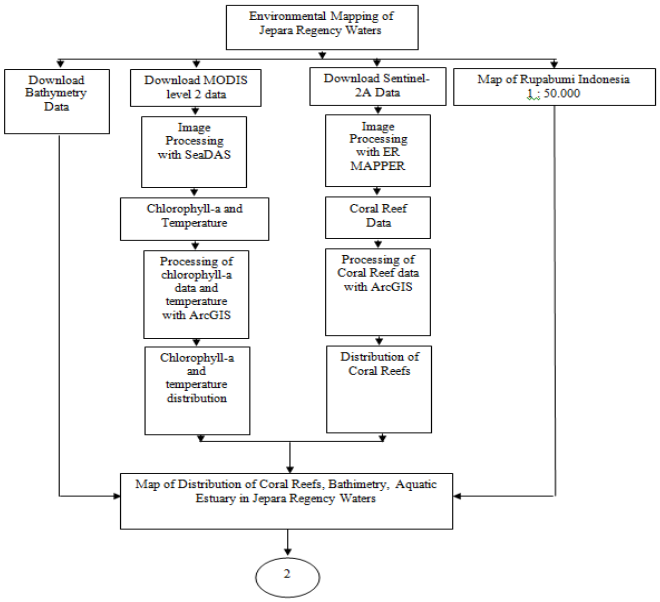


Figure 3: Flow chart of the distribution of vital habitat (coral, estuarine) and environmental conditions on the coast of Jepara Regency

- The third stage

The third stage is mapping the potential area of demersal fish. The third step is verification of field corals data and note the local name of the corals. All data both image data and field data are overlaid so formed maps of corals habitat distribution data and oceanographic factors, including bathymetry and TSS. Subsequently, an analysis was conducted to plot the potential catchment area of demersal fisheries (red snapper, coral grouper, sunu grouper, stingray, and lobster). Plotting of fish potential area based on the environmental

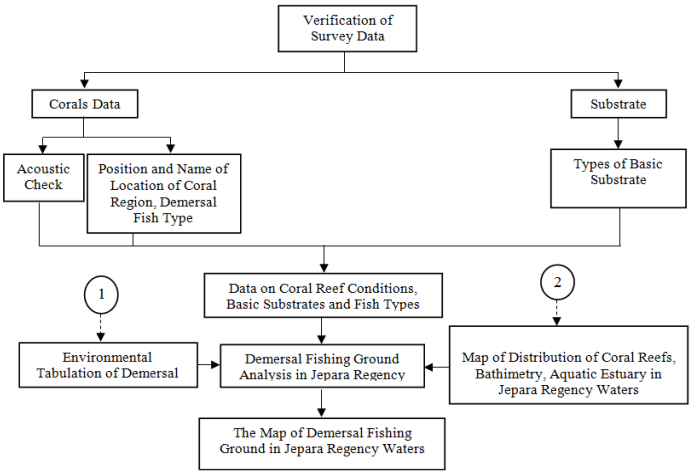


Figure 4. Flow chart prediction of demersal fishing ground mapping on the coast of Jepara Regency

3. Result and Discussion

3.1. The Economic Demersal Fish Species Associated with Corals in Jepara Coastal Waters

Based on the reference study and a survey at a fishing landing port in Jepara there are 10 economical demersal fish from Jepara Beach were found, namely Sea Catfishes, Red Snappers, Blue Line Seabass , Leopard Coral Grouper, Hairtails, Rays,

Lobster, White Shrimp, Rainbow Shrimp and Swimming Crabs (Figure 5). Based on the data there are only 4 types of economic demersal fisheries associated with corals were obtained from Jepara coastal waters, namely: Lobster, Red Snappers, Blue Line Seabass and Leopard Coral Grouper. Based on the environmental types of potential demersal fisheries in Jepara Regency Waters shown in Table 1, Lobster from childhood to adulthood live in a coral environment, a childhood in the bottom of sandy waters and an adult verada in the sidelines or hole of a coral [9]. Red Snappers and, Blue Line Seabass of his childhood life in beaches and river estuaries, but when it has begun to mature migrate to the rocky area [10]. Leopard Coral Grouper in childhood spawn and develop in seagrass areas, but when they mature they migrate to coral reef areas [10]. Demersal fisheries consisting of fish and crustaceans associated with corals are mature and foraging in coral areas. They are caught with fishing gear and traps in the coral reef area. However, when approaching spawning the corals fish migrate to the seagrass, muddy or sandy coastal areas.

Table 1: Environmental Conditions of Economical Demersal Fisheries Fish in Jepara Regency Table

Num	Type of Habitat/ Associated fish types						
	Corralling Beach/ Rocky	Seagrass Beach	Mangrove Beach	Estuarine	Offshore	Substrate	Depth/ Optimum (meters)
1	Lobster (larvae-adults					Corals-sand / mud- sand	<100/7-15
2			Sea Catfishes (larvae - young)	Sea Catfishes (adult)		mud	<100
3	Red Snappers (adult)			Red Snappers (larvae - young)		Mud mixed with coral fragments	<100/40-70
4	Blue Line Seabass (adult)			Blue Line Seabass (larvae - young)			<40
5	Leopard Coral Grouper (adult)	Leopard Coral Grouper (larvae-young)					<40/7-9
6				Hairtails (lavae)	Hairtails (adult)	mud	<2000/ 1-250
7	Rays			Rays		Sand-mud	<90/ <40

3.2. Prediction of Demersal Fishing Ground Associated with Corals in the Coastal Jepara Regency

Determination of demersal fishing ground is different from determining pelagic fishing ground. The fishing ground for demersal fish is closely related to the condition of its vital habitat, namely the location of corals, seagrasses, mangroves, and river mouths. Another limiting factor is the base substrate and the depth of the water. The dependence on these factors makes demersal fishing groundfish such as Red Snapper, Blue Line Seabass, Leopard Coral Grouper, and Lobster likely to remain location. So that our understanding of the habitat location of the demersal fisheries, makes it easier to determine the prediction of its fishing ground.

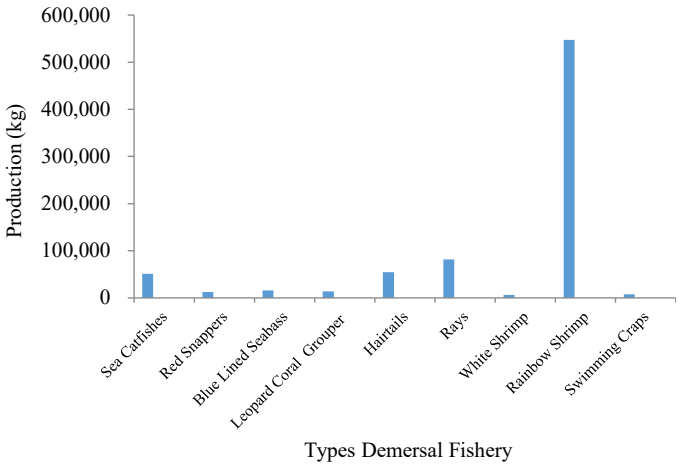


Figure 5. Demersal fisheries production in Jepara Regency (DKP Jepara, 2017)

Based on the results of the Sentinel-2 satellite image data processing and field survey, the Jepara Regency Coastal Area from West to East was divided into 7 segments (Figure 1), apparently only 4 segments contained coral reef habitats. The four segments are segment 2 namely the coastal waters of Tahunan and Jepara Districts (Figure 5), segment 3 namely the waters of the Districts of Jepara and Mlonggo (Figure 6), segment 4 namely the coastal waters of the Districts of Mlonggo and Bangri (Figure 7), and segment 5 which is the waters coast of Kembang and Keling Districts (Figure 8).

Three segments were not found coral reefs, that is in segment 1 namely Kedung District Coastal waters, segment 6 namely Keling and Donorojo District Coastal waters and segment 7 in the Donorojo District Coastal waters (Figure 1). Based on the field survey, the three segments that were not found in the coral reef had a substrate dominant with mud.

- *Prediction of Demersal Fishing Ground in Segment 2*

Segment 2, which covers the waters of the Tahunan District up to part of Jepara District based on Sentinel-2 satellite imagery and a field survey found 8 groups of coral reefs, with the local name and approximate area of Bokor Reef (5 ha), Teluk Awur Reef (12 ha), Tegal Sambi Reef (3 ha), Dalem Reef (3 ha), Pemandian Reef and Pulau Reef (66.4 ha), Panjang Island Reef (8 ha) and Sekumbu Reef (17.5 ha). Based on Sentinel-2 imagery, all clusters of reefs can be identified in terms of area and position, except for Bokor Reef only identified with field survey by snorkeling.

Seven other coral reef visible from Sentinel imagery are predicted as demersal fishing ground, especially type of fish Red Snappers, Blue Line Seabass and Leopard Coral Grouper, in green clusters (Figure 5). The highest potential demersal fishing ground

in terms of the extent of coral reef is predicted to be in the east coast region of Jepara District, which is along the location between Pemandian Reef and Sekumbu Reef.

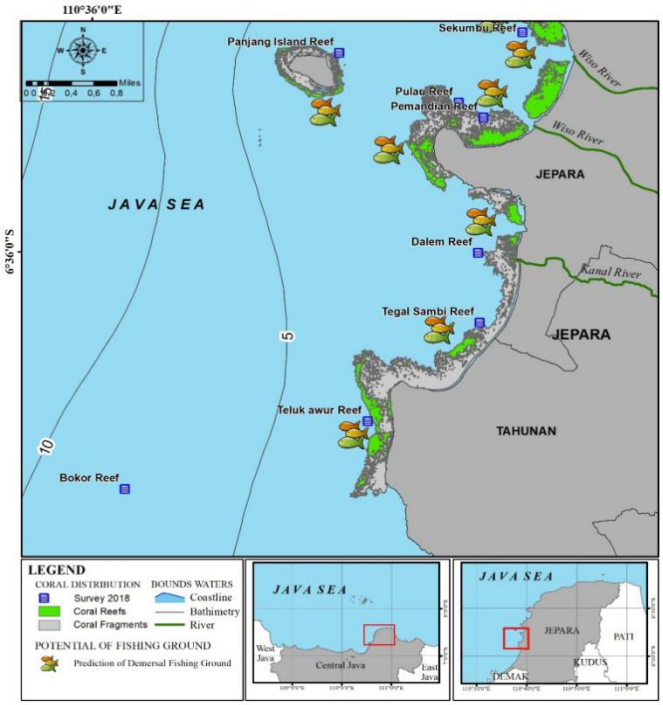


Figure 5. Predictions of potential fishing ground associated with corals in Coastal of Jepara segment 2 (Tahunan and Jepara Districts)

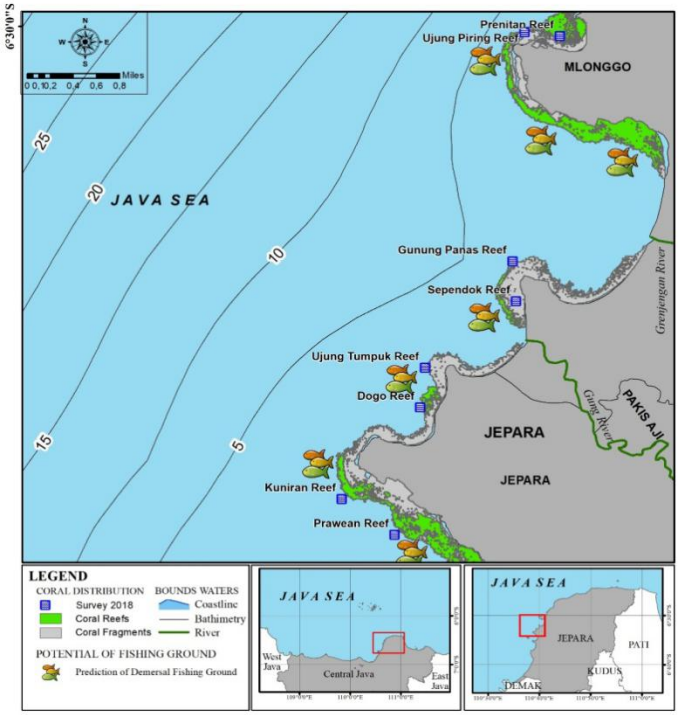


Figure 6. Predictions of potential fishing ground associated with corals in Coastal of Jepara segment 3 (Jepara and Mlonggo Districts)

- *Prediction of Demersal Fishing Ground in Segment 3*

Segment 3, which covers the waters of the Jepara District up to part of Mlonggo District based on Sentinel-2 satellite imagery and a field survey found 8 groups of coral reefs, with the local name and approximate area of Prawean Reef (17.5 ha), Kuniran Reef (38 ha), Dogo Reef (10 ha), Ujung Tumpuk Reef (14 ha), Sependok Reef (20 ha), Gunung Panas Reef (9 ha), Ujung Piring Reef (39 ha) and Prenitan Reef (9 ha) (Figure 6).

Six coral reef visible from Sentinel imagery are predicted as demersal fishing ground, especially type of Red Snappers, Blue Line Seabass and Leopard Coral Grouper, and lobster in green clusters (Figure 6). The Six coral reef are Prawean Reef, Kuniran Reef, Dogo Reef, Sependok Reef, Ujung Piring Reef, and Prenitan Reef. The highest potential demersal fishing ground in terms of the extent of coral reef is predicted to be in the coast region of Mlonggo District, which is along the location between Ujung Piring Reef than followed Prawean Reef.

- *Prediction of Demersal Fishing Ground in Segment 4*

Segment 4, which covers the waters of the Mlonggo District up to part of Bangsri District based on Sentinel-2 satellite imagery and a field survey found 9 groups of coral reefs, with the local name and approximate area of Semayit Reef (2 ha), Sejongok Reef (5 ha), Purancak Reef (5 ha), Kemangi Reef (5 ha), Ombo Reef (19.6 ha), Pandan Reef (2 ha), Lasak Reef (9 ha) and Ngelo Utara Reef (dead) dan Selaran Reef (dead) (Figure 7).

Four coral reefs visible from Sentinel imagery are predicted as demersal fishing ground, especially type of Red Snappers, Blue Line Seabass and Leopard Coral Grouper, and lobster in green clusters (Figure 7). The Four coral reef are Sejengok Reef, Purancak Reef, Ombo Reef, and Lasak Reef. The highest potential demersal fishing ground in terms of the extent of coral reef is predicted to be in the coast region of Ombo District, which is in offshore depth 10 m than followed Lasak Reef.

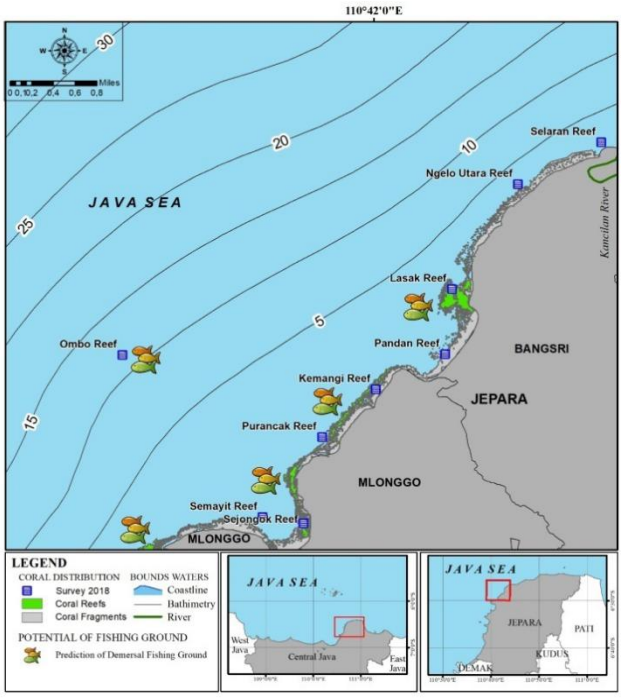


Figure 7. Predictions of potential fishing ground associated with corals in Coastal of Jepara segment 4 (Mlonggo and Bangsri Districts)

- *Prediction of Demersal Fishing Ground in Segment 5*

Segment 5, which covers the waters of the Kembang District up to part of Keling District based on Sentinel-2 satellite imagery and a field survey found only one group of coral reefs, with the local name and approximate area is Segedek Reef (10 ha) (Figure 8). Segedek Coral reef is about 4 m under surface waters, it's location in the Coastal of Keling District. That coral reefs are found by field survey predicted as demersal fishing ground, especially the type of Red Snappers, Blue Line Seabass and Leopard Coral Grouper, and lobster (Figure 8).

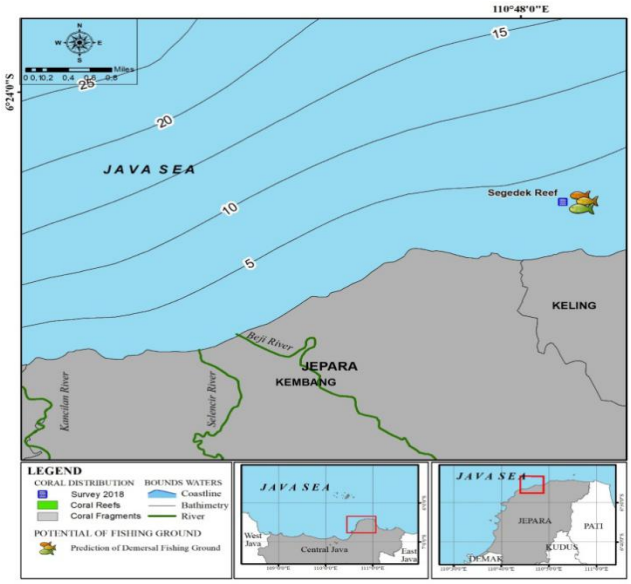


Figure 8. Predictions of potential fishing ground associated with corals in Coastal of Jepara segment 5 (Kembang and Keling Districts)

- *Limitations on the use of Sentinel-2 Satellite Imagery to be applied to Corals*

The use of Sentinel-2 satellite imagery to analyze the distribution of coral reefs in the coastal area of Jepara Regency turned out to be a problem so that some areas were not visible. Two things that are suspected as the main cause are spatial resolution and water turbidity.

3.3. Spatial Resolution

The sentinel image is used sentinel 2A, with a spatial resolution of 10 m. This pixel is influenced by the environment inside the pixel itself and outside the pixel. If the pixel is dominated by sand with only a few corals, then the coral will be invisible. Likewise, the environment around the pixel also affects, if around the pixel it turns out to be dominated by sand, while appears is sand. So that one pixel of coral can be seen the coral then there is a minimum of 3 x 3 pixels covered by coral. This is in accordance with what explained, which states that pixels outside the sample area need to be considered, and related to the size of the object (too small) will be a difficulty. So for the use of sentinel 2A imagery, the minimum coral cover that can be identified is 30 m^2. Examples of corals in Jepara that are present but not visible in the image are Karang Bokor on the west coast of Teluk Awur, Karang Ombo, Karang Mandalika and Karang Beteng (Figures 5, 7 and 9).

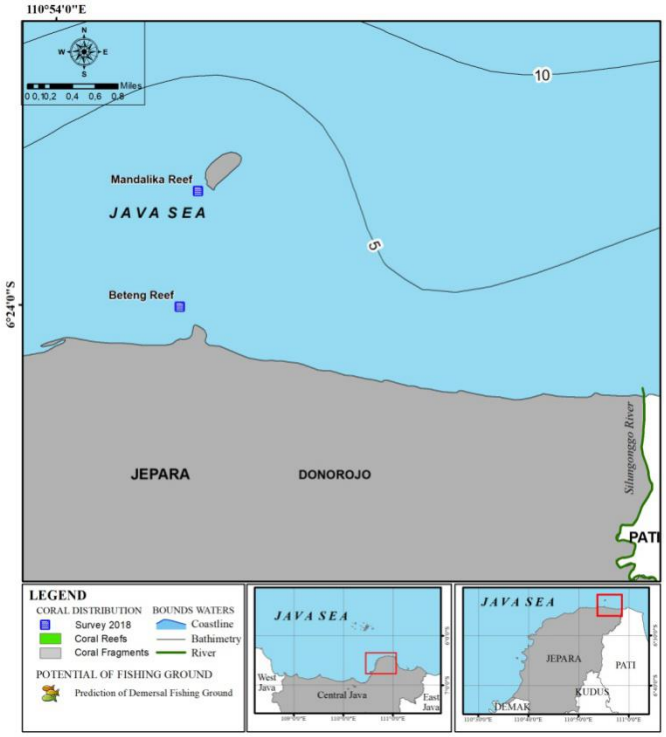


Figure 9. Beteng Reef and Mandalika Reef based on field surveys found its location in the northern waters of the Donorojo Regency, but from Sentinel Citra-2 it was not seen.

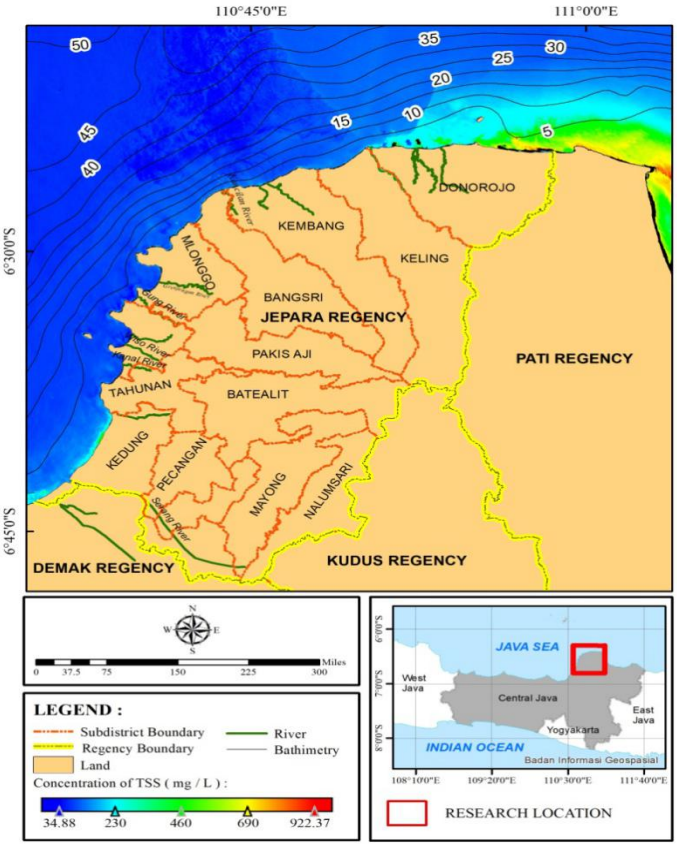


Figure 10. The Coastal Waters Northern Donorojo District and Western Kedung District contain high turbidity because high TSS (in the red squares)

3.4. Water Turbidity

Water turbidity is a serious obstacle. Coral reefs habitats in the northern coastal region, starting from Keling, Kembang, Donorojo District Waters with identification using Sentinel 2A images, did not show any coral. This is thought to be related to the high turbidity value, which is caused by the flow of rivers carrying mud from the Pati District area. Water masses carrying high Total Suspended Solids (TSS) values with range concentration 250-690 mg/L (Figure 10) are carried by ocean currents entering the Jepara region (Figure 10). This high TSS mass makes the sun's electromagnetic energy unable to penetrate the water so that the coral objects are not visible. Presumably, the reflection of the suspended solids, makes the sea surface closed, and looks like sand (Figure 10).

4. Conclusion

Innovation of demersal fishing ground prediction can be done in three stages, fist tabulating demersal fisheries environmental conditions, second mapping of vital habitat (coral reefs) and its supporting parameters, third mapping prediction of the potential area of demersal fish. Namely, demersal fishing ground can be made based on an understanding of childhood fish habitat, adulthood, substrate base habitat, an optimum depth of catchment area. Sentinel satellite imagery can be used to help determine the fishing ground of demersal fish species, but for the coastal waters of Jepara Regency, there are still obstacles related to spatial resolution and water turbidity. Field data is still needed to help determine the demersal fishing ground. Based on the fishing ground mapping, there are four economical types of fish that live in the coral reef environment, namely lobster, red snapper, coral grouper, and tiger grouper. The demersal fishing ground associated with corals in the Coastal Jepara Regency Waters which is seven segments apparently only four segments have the potential to become demersal fishing ground associated with corals, that are the second to fourth segments, including the coastal waters of Jepara, Mlonggo, Bangsri and Kembang District and the most potential is the coastal waters of Mlonggo District.

Acknowledgment

We thank the Diponegoro University for receipt of research funds with contract number 1501-33 / UN7.5.10 / LT / 2018 and Head of the Oceanography Department, Dr. Denny Nugroho S, ST., M.Sc, for the support of easy research administration. We also thank two of my assistants namely Mr. Bayu Munandar and Mr.Dika Ahmad Rojikin for their help so far.

References

[1] Kunarso, "Opportunities for Improvement of Fish Potential Location Forecast Models (Development of New Models for Increasing the Accuracy and Coverage of Fishing Area", Paper of Focus Group Discussion in Topics "Maximizing Fish Potential Information for Fishermen" in the forum ICCTF (Indonesian Climate Change Trust Fund) and National Planning Agency of the Republic of Indonesia (BAPENAS), Jakarta June 30th 2014, 16 pages, 2014.

[2] Lalli, C.M. and Parson, T.R., "Biological Oceanography: An introduction", Pergamon, BPC Wheatons Ltd, British.1994

[3] DKP Jepara (Jepara Regency Fisheries Service), "Profile Book of Fisheries Sector in Jepara Regency", Fisheries Service, Jepara Regency Government-Indonesia, 51 pages, 2017.

[4] Hedley, J.D., Roelfsemab, C., Brandoc, V., Giardinoc, C., Kutserd, T., Phinnb, S., Mumbye, P.J., Barrilerof, O., Laporteg, J., Koetzh, B., "Coral reef applications of Sentinel-2: Coverage, characteristics, bathymetry and benthic mapping with comparison to Landsat 8", Remote Sensing of Environment. Vol. 216: 598-614 pp, 2018. https://doi.org/10.1016/j.rse.2018.07.014

[5] Jaelani L M, Laili N, and Marini. Y, "Pengaruh Algoritma Lyzenga dalam Pemetaan Terumbu Karang Menggunakan Worldview-2, Studi Kasus: Perairan PLTU Paiton Probolinggo". Research Report: Institusi Teknologi Sepuluh Nopember, Surabaya Indonesia, 2015.

[6] Gernez, P., Lafon, V., Lerouxel A., Curti C., Lubac, B., Cerisier,S., Barile' , L., "Towar Sentinel-2 High Resolution Remote Sensing of Suspended Particulate Metter in Very Turbit Waters: SPOT4 (Take5), Experiment in the Laore and Geronde Estuaries", Remote Sensing, 7, 9507-9528, 2015. DOI: 10.3390/rs 70809507.

[7] Bioresita F., Pribadi, C.B., Firdaus, H.S., Hariyanto, T., Puissant, A., "The Use of Sentinel-2 Imagery to Total Suspended Solids (TSS) Estimation in Porong River", Sidoarjo, Elipsoida Vol. 01, No 01, 2018.

[8] Green, E., Edwards, A. and Mumby, P., "Mapping bathymetry" In A.J. Edwards (Ed.), Remote sensing handbook for tropical coastal management (pp.219-234). Paris: UNESCO, 2000.

[9] Holthuis, L.B.," FAO Species Catalogue. Vol. 13. Marine Lobsters of the World", An annotated and illustrated catalog of species of interest to fisheries known to date, *FAO Fisheries Synopsis,* No. 125, Vol. 13. Rome, FAO. 1991. 292 p, 1991.

[10] Habibi, A., Sugianta, Yusuf, C., "Grouper and Snapper Fish- Guidelines for Catching and Handling", Word Wildlife Fund (WWF) Indonesia, 34 pages, 2011.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 267-273 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Enabling 3D Heterogeneous Structures Towards Smart Chips: A Review

Cheng Li, Mengfu Di, Zijin Pan, Albert Wang*

Department of Electrical and Computer Engineering, University of California, Riverside, CA 92521, USA

ARTICLE INFO

Article history:
Received: 23 November, 2019
Accepted: 19 January, 2019
Online: 27 January, 2020

Keywords:
Magnetic-cored inductor, nano crossbar array, graphene, NEM, switch, ESD protection, BEOL, metal wall, noise, non-binary

ABSTRACT

This review paper discusses recent research outcomes of fabricating three-dimensional (3D) heterogeneous structures in complementary metal-oxide-semiconductor (CMOS) integrated circuit (IC) technologies to enable smart future chips. The CMOS-based 3D heterogeneous structures demonstrated are vertical magnetic-cored inductors for radio-frequency (RF) ICs, through-back-end-of-line (BEOL) metal wall structures for global flying noise isolation, above-IC nano crossbar array electrostatic discharge (ESD) protection devices and graphene-based nano electromechanical (NEM) switch ESD protection structures, and individually controlled multiple-gated MOS field-effect transistors (MOSFET) for non-binary computing. These disruptive 3D heterogeneous structure concepts can potentially enable future smart IC chips in Si CMOS IC platforms.

1. Background

Since the birth of silicon monolithic integrated circuits (ICs) in 1959, geometrical scaling has been driving advances in silicon IC technologies, keeping on increasing transistor count at the Moore's Law pace until late 1990s [1]. The hallmark Si planar process was the technical foundation that has been guided by the National Technology Roadmap for Semiconductor (NTRS) through 1990s [2]. The laterally dimensional scaling has evolved into equivalent scaling, as guided by the International Technology Roadmap for Semiconductor (ITRS) until 2000 [3], where new materials and structural features, such as strained silicon and FinFET, were included. Later, 3D Power scaling kicked in to accommodate the ever-increasing demands for and emerging challenges of density, complexity and power as IC technologies continue to advance to sub-10nm nodes, as guided by ITRS 2.0 until 2015 [4]. Nevertheless, as 7nm CMOS is already in production today and 3nm CMOS is on the horizon, lateral shrinking is quickly approaching its end. Since the horizontal Si "land" is limited for increasingly complex ICs, going vertical to make Manhattan style Si-based "skyscrapers" becomes an obvious alternative. Meanwhile, future chips are not only bigger and more complicated, they ought to be smarter too, demanding for more device functionalities to handle more information diversity. Consequently, the ITRS 2.0 was brought to its end in 2015 [5] and the new Heterogeneous Integration Roadmap (HIR) was initiated to address the emerging technical needs and challenges in the new era of integrated circuits and systems [6-7]. The required functional diversification cannot not be delivered by conventional Si IC devices. Many novel devices and structures will be needed, which must be heterogeneously integrated into Si ICs in a 3D fashion, mainly based upon the CMOS platforms. To this end, substantial R&D efforts have been devoted to developing non-traditional 3D heterogeneous devices and structures to enhance functionality and improve performing characteristics of conventional Si-based CMOS ICs. For example, Nanoelectromechanical (NEM) relay structure in CMOS backend was recently reported to realize reconfigurable interconnects to enable complex and hard-reconfigurable ICs [8]. This review paper, an extended version of our paper presented at IEEE International Conference on IC Design and Technology [9], provides an overview of our recent research outcomes of developing various disruptive 3D heterogeneous structures to enable smart future CMOS chips.

2. Vertical Magnetic-Cored Inductors

Proliferation of wireless communications have changed our society and the foundation is RF IC technology. The ultimate design goal is to achieve RF system-on-a-chip (SoC) to ensure higher system performance and reliability while reducing costs. The main technical barrier for making RF SoCs is lack of super compact, multiple-GHz, high-performance IC inductors in CMOS. Traditional IC inductors using metal spirals are too large, typically bigger than a bonding pad, which makes single-chip RF SoC impractical. The dream for RF IC designers is to have high-

*Corresponding Author: Albert Wang, Email: aw@ece.ucr.edu

performance transistor-sized inductors in CMOS. For years, various magnetic-enhanced inductor structures have been reported, utilizing lateral magnetic films or complex MEMS structures to reduce magnetic losses. These magnetic inductors are impractical to CMOS ICs. As a solution, we proposed a novel stacked-via vertical-magnetic-cored inductor structure made in the back-end-of-line (BEOL) of CMOS [10-17]. To understand our disruptive concept, imagine to shrink a perfect discrete solenoid inductor and plugged it vertically into the CMOS BEOL stack, like a "needle", which forms a transistor-sized vertical-magnetic-cored inductor in CMOS BEOL stacks. As the inductor size shrinks, electrical loss increases due to increased series resistance of metal wires, which however, can be compensated by reducing magnetic loss by use of a closed-loop magnetic circuit. Hence, the overall inductor performance does not suffer from inductor shrinking. Figure 1 depicts the concept of the new vertical- magnetic-cored inductor. Taking full use of CMOS BEOL structures, the vertical magnetic bar can be fabricated by locally replacing the common metal vias in a CMOS BEOL stack with suitable magnetic materials, hence called a stacked-via magnetic-cored inductor. One critical specification for the magnetic-cored inductor is to achieve high performance (e.g., Q-factor and inductance density, L-density) at high frequencies (>1GHz). Substantial theoretical and experimental work has been conducted to prototype the new inductor devices and improve the inductor performance [10-19]. Various magnetic materials compositions, materials synthesis techniques, magnetic core structures and process steps were explored. In our initial work, ferrite thin film materials, e.g., Ni0.4Zn0.4Cu0.2Fe2O4, Y2.8Bi0.2Fe5O12 and Co7ZrO9, were used as the magnetic media. A CMOS-compatible process module was developed using a spin-coating method to make single-layer ferrite magnetic enhanced inductors. As depicted in Figure 2, measurements showed encouraging inductance improvement over the non-magnetic references (R1/R2), from 5.8% to 38% up to 9GHz. To enhance magnetic materials for multiple-GHz operations, we further engineered various nano particle powders for the magnetic cores, including Co_2Z-type ($Ba_3Co_2Fe_{24}O_{41}$) and Ni–Zn–Cu ($Ni_{1-x-y}Zn_xCu_yFe_2O_4$) series. We found that the Ni–Zn–Cu ($Ni_{1-x-y}Zn_xCu_yFe_2O_4$) composites achieved the best reported performance up to ~10GHz. We then developed a post-CMOS process flow to fabricate the vertical magnetic cores for RF inductors as shown in Figure 3. The new vertical magnetic-cored inductors were fabricated in a foundry 180nm 6-metal CMOS technology. Shown in Figure 4, the new magnetic-cored inductors achieved the then- best-reported inductor performance, i.e., Q_{max}·f_{max}~L-density characteristics: ~220% improvement in Q-factor to up to 20GHz and the highest reported L-density of ~700nH/mm^2 at $Q_{max}f_{max}$ of ~4GHz. A new 3D modeling technique was developed to accurately simulate any arbitrary conductor-magnet structures to guide the design and optimization of vertical magnetic-cored inductors [18, 19]. The new stacked-via magnetic-cored structure can be used to fabricate other types of compact high-performance inductive devices, which may not only reduce RF IC chip size, but also potentially allow new chip architecture. For example, a bank of compact magnetic-cored transformers can be used to realize simplified power amplifier (PA) architecture to address the power back-off problem. The new magnetic-cored inductors were applied to design of LC-VCO ICs in CMOS [20, 21]. Figure 5 shows that in a 1.36–1.86GHz VCO fabricated in a 180nm RF CMOS, it achieves a high FOM factor of ~202dBc/Hz and a reduced phase noises of -121 dBc/Hz at 100-kHz [12]. In a 2.22-2.92GHz VCO fabricated in a 180nm SOI CMOS, it achieves phase noise reduction for from -106.97dBc/Hz to -113.49dBc/Hz [21].

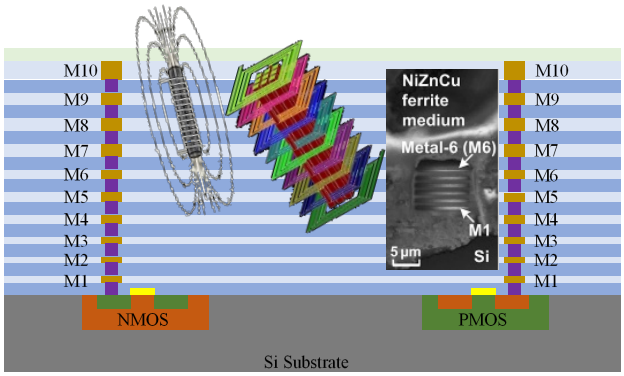


Figure 1 A conceptual view of new vertical magnetic cored inductor in CMOS platform. Insets from Left to Right: a discrete solenoid, simulated vertical magnetic-cored inductor and fabricated vertical magnetic-cored inductor.

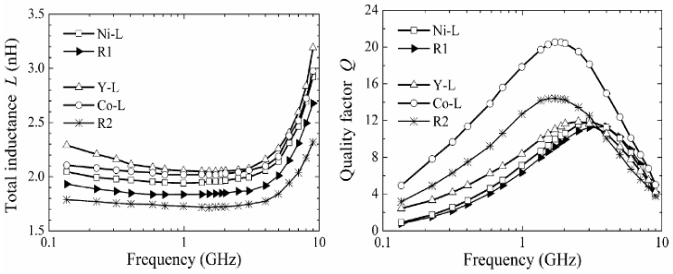


Figure 2 Measurement of initial single-layer magnetic-enhanced inductor shows improvement in inductance and Q-factor: Ni-L *versus* R1 and Y-L/Co-L *versus* R2 [10].

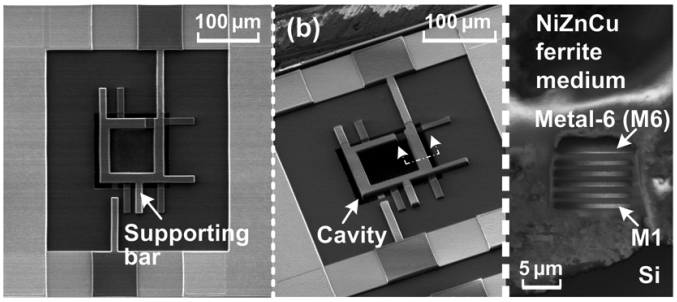


Figure 3 SEM photos for samples of fabricated vertical magnetic-cored inductors in a 180nm 6-metal CMOS.

3. Through-BEOL Metal Wall Noise Isolation

Interferences, i.e., noise or crosstalk, affect IC performance. As technology shrinking and chip scale increasing continue, noise/crosstalk isolation becomes extremely important and challenging, particularly for large and complex SoC chips. For years, substantial research has been devoted to develop various noise-isolation techniques, resulted in many isolation methods, such as, guide rings, deep trenches, etc. These noise-isolation structures are all inside Si substrates, hence, they only work for in-Si noise isolation.

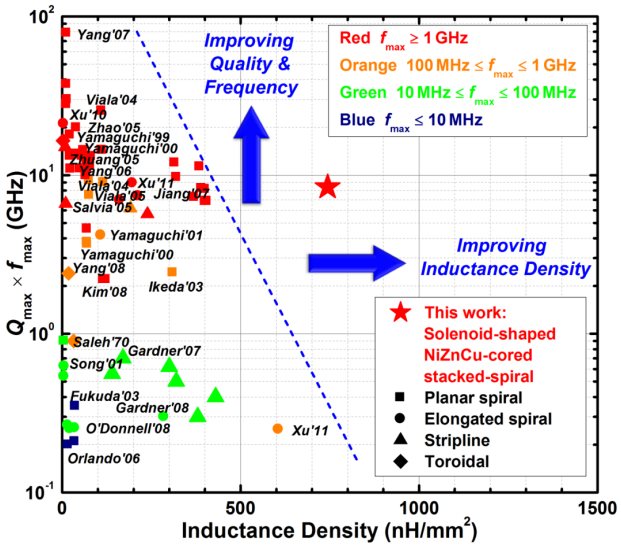


Figure 4 Comparison shows that the vertical magnetic-cored inductor (the star) out-performs the reported state-of-the-art of various magnetic-enhanced inductors.

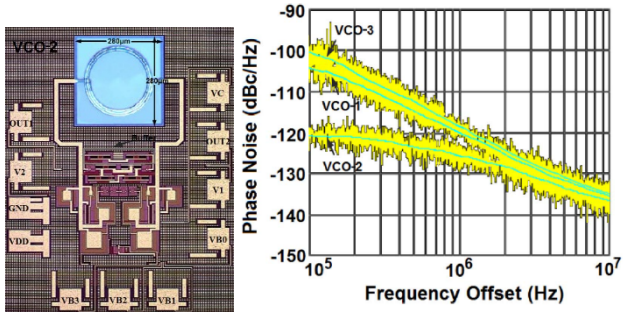


Figure 5 Measurement shows that VCO2 using magnetic-cored inductor has better phase noise performance compared to VCO1/VCO3 using conventional inductors [20].

For complex ICs at advanced nodes, noise coupling mainly occurs in the BEOL stacks through the extremely complex and conductive metal interconnects mesh. This new type of in-BEOL noise coupling is defined as "flying noise", which is global in nature across the whole chip, and unfortunately, cannot be suppressed by any conventional in-Si noise-isolation techniques. To address this emerging global flying noise challenge, we proposed a novel through-BEOL metal wall flying noise isolation structure as illustrated in Figure 6 [22, 23]. In principle, it is a (near) closed-loop metal wall structure that blocks transmission of electromagnetic waves, hence, noise-isolates circuit blocks residing inside the through-BEOL metal cages. From EM theory, the through-BEOL meta walls can be optimized in terms of metal materials and wall physical designs to block the noise propagation, therefore, eliminating the global flying crosstalk in the BEOL stacks. This flying noise isolation concept was validated experimentally in CMOS. Figure 7 shows the first experiment where a simple amplifier circuit was designed in a foundry 180nm SOI CMOS. A post-CMOS process module was developed to etch a deep trench into the CMOS BEOL stack, which was filled with silver powder to form the in-BEOL metal wall structure. In measurement, it was observed that the troublesome third-order intermodulation interferers were suppressed by about 9dB (Figure 7) [22].

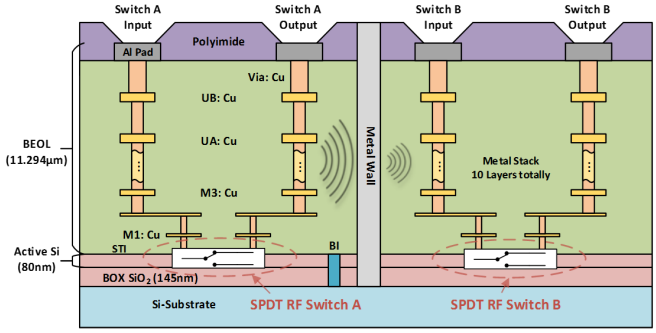


Figure 6 A conceptual X-section view of the new through-BEOL metal wall flying crosstalk isolation structure in SOI CMOS.

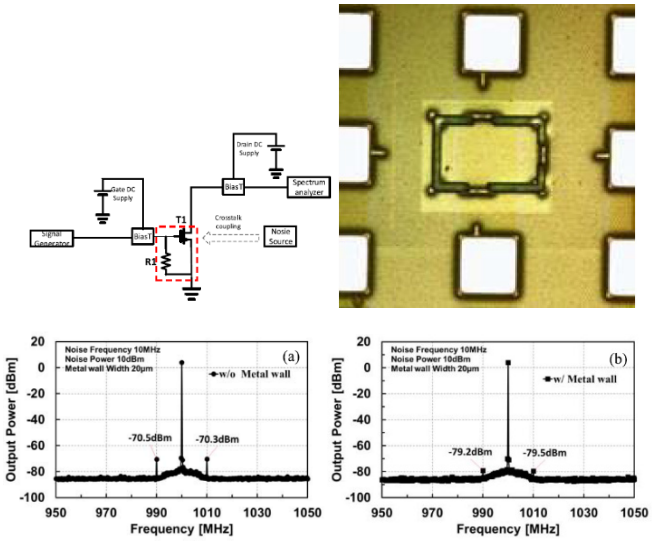


Figure 7 An amplifier circuit with the new in-BEOL metal wall can reduce the IM3 interferers by 9dB in measurement [22].

Figure 8 presents a second prototype design, which are SPDT RF switch circuits fabricated in a foundry 45nm SOI CMOS [23]. The design splits include SPDT switches with and without in-BEOL metal walls that were made of silver nano powders.

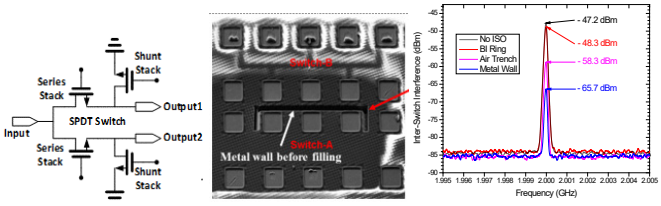


Figure 8 SPDT RF switches fabricated in a 45nm SOI CMOS utilizes the new through-BEOL metal wall to successfully block the global flying crosstalk by 98% in linear scale [23].

Measurement shows that the in-BEOL metal wall structure reduces the inter-switch flying crosstalk by 18.5dB, almost entirely blocked the global flying crosstalk (~98.6% in linear scale). The experiments confirmed that our new through-BEOL metal wall flying noise isolation technique is efficient in blocking the above-Si global flying crosstalk.

4. Above-IC Nano-Crossbar ESD Protection

Electrostatic discharge (ESD) failure is the most devastating reliability problem to ICs. On-chip ESD protection is required for all ICs and electronic systems [24]. As IC process technologies continue scaling and chip complexity rapidly increases, on-chip ESD protection design emerges as a major IC reliability design challenge. Since the initial ESD awareness in late 1960s, ESD protection has been relied on in-Si PN-junction-based active devices to form low-impedance conduction channels to discharge fast ESD transients, hence, provide ESD protection for ICs. Unfortunately, the conventional ESD protection structures inevitably introduce parasitic effects inherent to the PN junctions, such as ESD-induced parasitic capacitance, noises and leakage current, which seriously affect IC performance, particularly for high-frequency broad band RF ICs and high-speed ICs. On the other hand, on-chip ESD protection structures consume large Si die area in order to handle the significant ESD-related energy, and the typically irregular ESD device shapes makes IC physical layout design truly challenging. These ESD protection design problems are becoming unbearable to advanced ICs at sub-28nm nodes. It is hence imperative to develop disruptive ESD protection solutions for future chips that are not tied to in-Si PN junctions [25]. To address such emerging ESD protection design challenges, we devised two revolutionary above-IC non-PN type ESD protection concepts and device structures. Figure 9 depicts the first concept of an above-IC nano-crossbar array ESD protection structure [26-28]. In principle, this new device is a two-terminal phase-changing switch structure comprising two electrodes and a phase-changing insulator in between. The insulator is made of specially synthesized nano materials that is electrically OFF in normal IC operations. Under ESD stressing, the strong incident electrostatic force causes phase change of the nano insulator, making it conducting, hence turns ON the nano crossbar switch structure to discharge the ESD transients. When an ESD event is over, the switch will return to its OFF state. The nano crossbar array switch

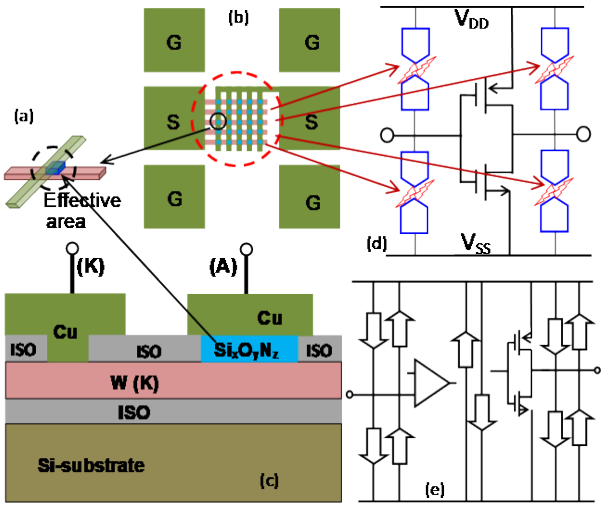


Figure 9 A conceptual illustration for the new above-IC nano crossbar array ESD protection structure in CMOS.

is built in the CMOS BEOL stack above the Si substrate and is not a conventional PN junction device. It is an above-IC ESD switch structure. While each crossbar node is a nano device, a large nano crossbar array can be designed to handle large ESD energy collectively. This novel nano crossbar array ESD switch structure was experimentally validated using a post-CMOS process module developed. Figure 10 presents the measured transient ESD I-V curve for a single-node crossbar ESD switch by transmission-line pulsing (TLP) ESD testing, showing the desired symmetrical two-directional ESD discharging I-V behavior. Figure 11 shows the measured I-C curves for a nano crossbar array ESD switch where multiple ESD triggering effect can be observed. Measurement shows ultrafast ESD response time of ~100pS, negligible leakage current of I_{leak}~2pA (Figure 12), tunable ESD triggering voltage V_{t1} of 1.26~28.3V (Figure 13) and ultrahigh ESD protection capability of ~267V/μm^2. Since the new nano crossbar ESD protection structure is above the Si substrate, it does not consume extra Si die area. This novel above-IC nano crossbar array ESD switch structure can potentially resolve all challenging problems associated with the conventional in-Si PN-based ESD protection structures.

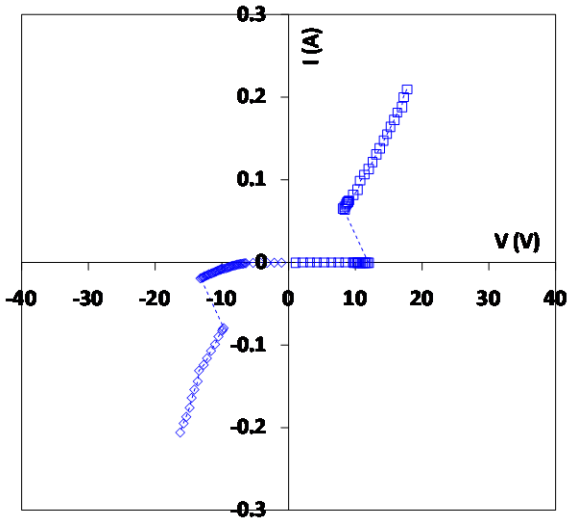


Figure 10 TLP testing shows symmetric ESD discharging I-V curve for a single-node nano crossbar ESD protection structure.

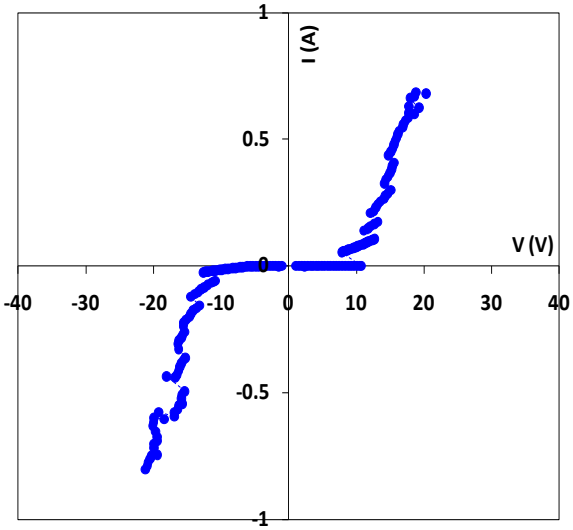


Figure 11 TLP testing shows symmetric ESD discharging I-V curve for a sample nano crossbar array ESD protection structure.

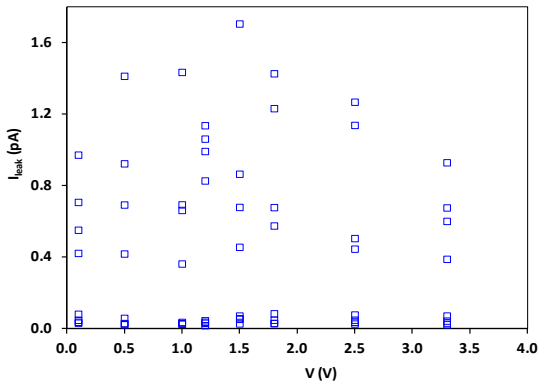


Figure 12 Measurement shows ultralow leakage currents, due to the non-PN ESD structure, for sample nano crossbar ESD protection structures.

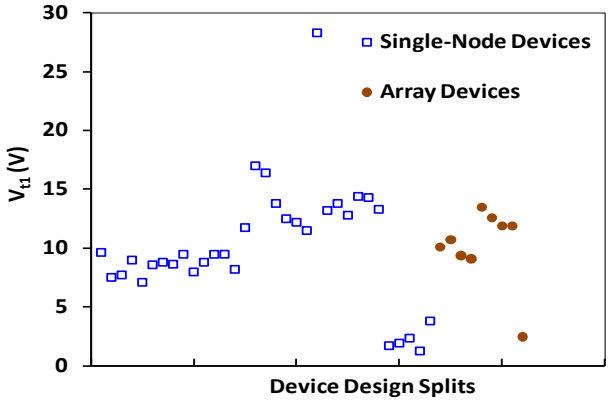


Figure 13 Measurement shows tunable ESD triggering voltage for sample nano crossbar ESD protection structures.

5. Above-IC Graphene NEMS ESD Switch

Figure 14 shows the second novel above-IC ESD protection structure, which is a graphene-based NEMS (gNEMS) switch ESD protection device fabricated in the CMOS BEOL stack [29, 30]. This new graphene gNEMS ESD switch is a two-terminal device, entirely different from conventional in-Si PN-based ESD protection devices. A gNEMS has a graphene membrane over a cavity. The two electrodes are the graphene membrane on top (cathode, or, K) and the bottom of the cavity (anode, or, A). Unlike any PN-based electronic devices, the gNEMS device is a mechanical switch, which remains electrically OFF in normal IC operations. When an ESD transient appears, the electrostatic force pulls the graphene membrane downward to touch the bottom conductor, hence turns on the gNEMS device to discharge the ESD surge. When the ESD transient is over, the graphene film will bounce back and turns the gNEMS OFF again. Since gNEMS is a mechanical switch, ideally, it has no leakage current and almost no parasitic capacitance. Since it is built in the CMOS BEOL stack, above the IC substrate, gNEMS devices will not consume extra Si die area. Therefore, the gNEMS ESD switch is an ideal ESD protection structure. Since graphene film is ultralight and has very large Young's module, the gNEMS switching speed can be extremely fast. Figure 15 presents the measured transient ESD discharging I-V behaviors for a sample gNEMS switch, showing the desired dual-directional and symmetrical I-V characteristic. Measurement also shows negligible leakage current of ~3pA, superfast ESD response time of ~200ps and ultrahigh ESD protection capability of ~1500V/μm^2, which is many orders of magnitude higher than a silicon-controlled rectifier (SCR) ESD protection structure that is normally considered the most robust Si ESD protection structure. It is encouraging that the gNEMS ESD switch can be potentially ideal ESD protection solution for future ICs.

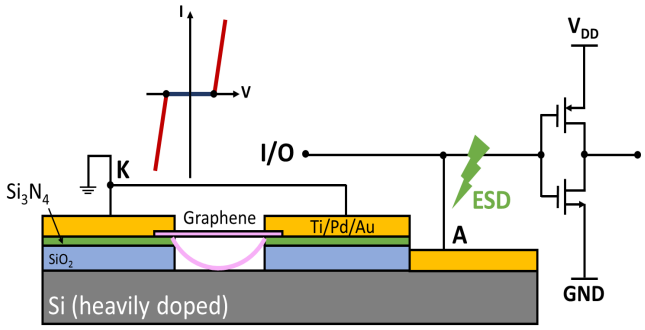


Figure 14 A conceptual illustration for the new above-IC gNEMS ESD switch device in CMOS.

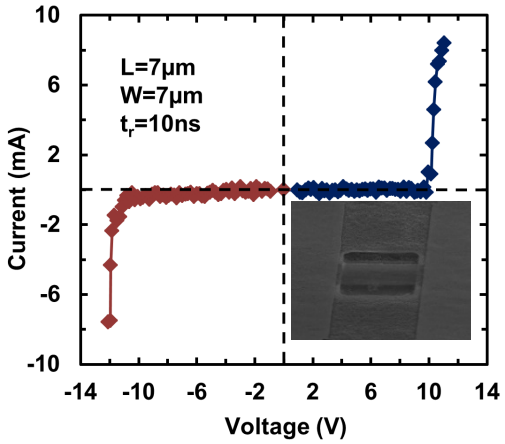


Figure 15 TLP testing shows a dual-directional ESD discharging I-V curve for a sample gNEMS ESD switch, where the inset is a SEM image for gNEMS.

6. Multiple-Gated MOSFETs

Our information and internet era rely on binary data processing, including computing and storage, which was enabled by transistors and CMOS ICs. The base-2 binary numbering system is the foundation for most information processing systems, whose capacity is certainly limited, i.e., 2^n (n is bit number in a byte) caps binary the information system capacity. The binary numbering system is natural to CMOS ICs because a MOSFET has two conduction states, OFF (0) and ON (1) corresponding to I_D = 0 or not. Imagine a base-M numbering system with M>2, where $M^n > 2^n$ results in a much bigger system capacity for a given byte size. Research has been done to explore multiple-state-value devices, mainly pursuing for unique conducting materials to realize many-status logics. As a potential solution, we proposed a novel individually-controlled multiple-gated MOSFET structure. Figure 16 depicts such a device concept where a MOSFET has four individually controlled gates over the same conduction channel. By selectively biasing the four gates, individually or collectively, this MOSFET will deliver multiple conduction states, hence, forms

a non-binary M^n numbering system [31, 32]. For different channel conduction states, I_D will be different for the MOSFET. By designing the 4-gated MOSFET, 5<M<16 may be obtained. There are several advantages for this base-M information MOSFET structure and numbering system. First, its information system capacity increases significantly from the binary system. Second, by choosing M=10 through device design, a base-10 (i.e., 10^n) system may be realized, which not only dramatically increases the system capacity, but also makes computing algorithm and system architecture more natural to human instinct. Third, this 4-gated MOSFET can be fabricated using classic CMOS IC technologies through 3D heterogeneous integration.

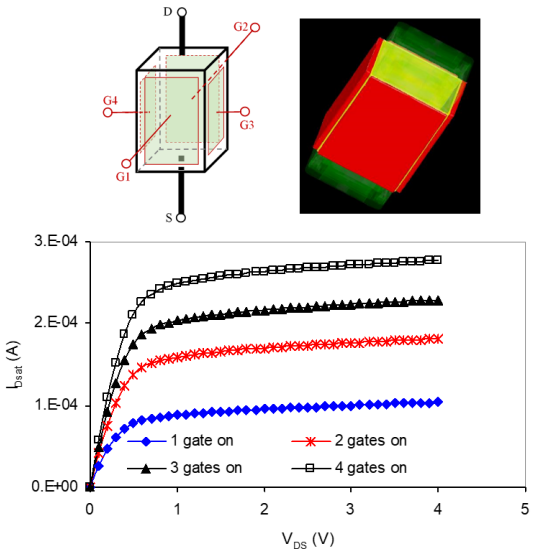


Figure 16 Conceptual illustration for the novel multiple-gated MOSFET.

7. Summary

The past six decades have witnessed the unprecedent advances of the Si-based IC technologies, which has completely changed human life. As we are rapidly moving from the information technology (IT) era to the internet-of-everything (IoET) age, it calls for disruptive IC technologies to facilitate the emerging needs for information sensing, computing, storage and communications, which simply cannot be supported by the hallmark Si planar IC process technologies. To break the deadlock of space, complexity and performance of ICs, one viable answer to the grand technical challenge is to develop smart future chips that can integrate more device functionalities, in a 3D fashion, to improve operating characteristics of IC chips. 3D heterogeneous structures and heterogeneous integration are the potential solutions to continuous advances of Si-based CMOS ICs beyond the Moore's Law. This paper reviews some novel 3D heterogeneous structures that we developed recently to address the emerging IC design challenges including: compact inductors with vertical magnetic cores, in-BEOL global flying noise isolation structures, on-chip ESD protection and multi-gated transistors for non-binary numbering systems. These new 3D heterogenous structure concepts, while being exemplar, are truly disruptive and potentially enablers for future smart chips.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors acknowledge National Science Foundation for partial supports of the related research.

References

[1] G. E. Moore, "Cramming more components onto integrated circuits", Electronics, pp. 114-117, April 19, 1965. DOI: 10.1109/N-SSC.2006.4785860
[2] SIA National Technology Roadmap for Semiconductor (NTRS)
[3] SIA International Technology Roadmap for Semiconductor (ITRS)
[4] SIA International Technology Roadmap for Semiconductor 2.0 (ITRS 2.0), 2015 Edition.
[5] ITRS wensite, http://www.itrs2.net/
[6] Heterogeneous Integration Roadmap (HIR), 2019 Edition.
[7] Heterogeneous Integration Roadmap (HIR) website, https://eps.ieee.org/technology/heterogeneous-integration-roadmap/2019-edition.html
[8] Z. A. Ye, S. Almeida, M. Rusch, A. Perlas, W. Zhang, U. Sikder, J. Jeon, V. Stojanović, and T.-J. K. Liu, "Demonstration of 50-mV Digital Integrated Circuits with Microelectromechanical Relays", in IEEE IEDM Tech Digest, 2018. DOI: 10.1109/IEDM.2018.8614663
[9] C. Li, F. Zhang, M. Di, Z. Pan and A. Wang, "Developing 3D Heterogeneous Structures for Future Chips", in Proc. IEEE International Conference on IC Design and Technology (ICICDT), pp.48-51, 2019. DOI: 10.1109/ICICDT.2019.8790942
[10] C. Yang, F. Liu, T. Ren, L. Liu and A. Wang, "On-Chip Integrated Inductors with Ferrite Thin-films for RF IC", in IEEE IEDM Digest, pp.225-228, 2006. DOI: 10.1109/IEDM.2006.347003
[11] J. Zhan, C. Yang, X. Wang, Q. Fang, Z.T. Shi, Y. Yang, T.-L. Ren, A. Wang, Y.H. Cheng, L.-T. Liu, "Stacked-spiral RF inductors with vertical nanoparticle magnetic core for radio-frequency integrated circuits in CMOS", Sensors & Actuators: A. Physical, V195, pp. 231-238, June 2013. http://dx.doi.org/10.1016/j.sna.2013.02.030
[12] Z. Ni, J. Zhan, Q. Fang, X. Wang, Z. Shi, Y. Yang, T.-L. Ren, A. Wang, Y. Cheng, J. Gao, X. Li, and C. Yang, "Design and Analysis of Vertical Nanoparticles- Magnetic-Cored Inductors for RF ICs", IEEE Trans. Electron Devices, **60**(4), pp. 1427-1435, April 2013. DOI: 10.1109/TED.2013.2245418
[13] J. Zhan, C. Yang, X. Wang, F. Zhang, T.L. Ren, A. Wang, Y. Yang, L. T. Liu, L. W. Yang and Z. X. Yue, "Stacked-Spiral RF Inductor with Vertical Nano-Powder-Magnetic-Core in CMOS", IEEE Microwave and Wireless Components Letters, **22**(1), pp.29-31, January 2012. DOI: 10.1109/LMWC.2011.2176474
[14] C. Yang, F. Liu, X. Wang, J. Zhan, A. Wang, T-L. Ren, L-T. Liu, H. Long, Z. Wu, X. Li, "Investigation of on-Chip Soft-Ferrite-Integrated Inductors for RF ICs — Part I: Design and Simulation", IEEE Trans. Electron Devices, **56**(12), pp.3133-3140, December 2009. DOI: 10.1109/TED.2009.2033328
[15] C. Yang, F. Liu, X. Wang, J. Zhan, A. Wang, T-L. Ren, L-T. Liu, H. Long, Z. Wu, X. Li,, "Investigation of on-Chip Soft-Ferrite-Integrated Inductors for RF ICs — Part II: Experiments", IEEE Trans. Electron Devices, **56**(12), pp. 3141-3148, December 2009. DOI: 10.1109/TED.2009.2033413
[16] C. Yang, F. Liu, T. Ren, L. Liu, G. Chen, X. Guan, A. Wang and H. Feng, "Ferrite-Integrated on-Chip Inductors for RF ICs ", IEEE Electron Device Letters, **28**(7), pp.652-655, July 2007. DOI: 10.1109/LED.2007.900299
[17] C. Yang, F. Liu, T-L. Ren, L-T. Liu, G. Chen, X-K. Guan, A. Wang and Z-X. Yue, "Ferrite-Partially-Filled on-Chip RF Inductor Fabricated Using Low-Temperature Nano-Powder-Mixed-Photoresist Filling Technique for Standard CMOS", in IEEE IEDM Tech. Digest, pp1038-1040, 2007. DOI: 10.1109/IEDM.2006.347003
[18] H. Long, Z. Feng, H. Feng and A. Wang, "A novel accurate PEEC-based 3D modeling technique for RF devices of arbitrary conductor-magnet structure", Microwave and Optical Technology Letters, 38(3), Wiley & Sons, pp237-240, August 2003.
[19] H. Long, Z. Feng, H. Feng, A. Wang, T. Ren, J. Bao, F. Liu, C. Yang, X. Zhang, "A New Modeling Technique for Simulating 3D Arbitrary Conductor-Magnet Structures for RFIC Applications", IEEE Tans. Electron Devices, **52**(7), pp. 1354-1363, July 2005. DOI: 10.1109/TED.2005.850662
[20] H. Cai, Y. Yang, N. Qi, X. Chen, H. Tian, Z. Song, Y. Xu, C-J. Zhou, J. Zhan, A. Wang, B. Chi and T-L. Ren, "A 2.7-mW 1.36–1.86-GHz LC-VCO With a FOM of 202 dBc/Hz Enabled by a 26%-Size-Reduced Nano-Particle-Magnetic-Enhanced Inductor", IEEE Trans. Microwave Theory and Techniques, **62**(5), pp. 1221-1228, May 2014. DOI: 10.1109/TMTT.2014.2312886
[21] R. Ma, F. Lu, Q. Chen, C. Wang, F. Liu , W. Zou and A. Wang, "A 2.22–2.92GHz LC-VCO Demonstrated with An Integrated Magnetic-Enhanced

Inductor in 180nm SOI CMOS", in Proc. IEEE RFIC Symp., pp.110-113, 2016. DOI: 10.1109/RFIC.2016.7508263

[22] C. Wang, Q. Chen, F. Lu, C. Li, F. Zhang and A. Wang, "Blocking Flying Crosstalk in BEOL Validated in Antenna Switches in 45nm SOI CMOS", IEEE Microwave and Wireless Components Letters, **28**(11), pp.1005-1007, November 2018. DOI: 10.1109/LMWC.2018.2873197

[23] F. Lu, Q. Chen, C. Wang, F. Zhang, C. Li, R. Ma, X. S. Wang and A. Wang, "In-Die Through-BEOL Metal Wall for Noise Isolation in 180nm FD-SOI CMOS", IEEE Electron Device Letters, **38**(5), pp.630-632, May 2017. DOI: 10.1109/LED.2017.2682819

[24] A. Wang, On-Chip ESD Protection for Integrated Circuits, Kluwer, Boston, 2002.

[25] A. Wang, H. Feng, R. Zhan, H. Xie, G. Chen, Q. Wu, X. Guan, Z. Wang and C. Zhang, "A review on RF ESD protection design", IEEE Trans. Electron Devices, **52**(7), pp.1304-1311, July 2005. DOI: 10.1109/TED.2005.850652

[26] L. Wang, X. Wang, Z. Shi, R. Ma, J. Liu, Z. Dong, C. Zhang, L. Lin, H. Zhao, L. Zhang, Y. Cheng, R. Huang and A. Wang, "Dual-Directional Nano Crossbar Array ESD Protection Structures", IEEE Electron Device Letters, **34**(1), pp.111-113, January 2013. DOI: 10.1109/LED.2012.2222337

[27] L. Lin, L. Zhang, X. Wang, J. Liu, H. Zhao, H. Tang, Q. Fang, Z. Shi, R. Huang, Y. Cheng and A. Wang,, "Novel Nanophase-Switching ESD Protection", IEEE Electron Device Letters, **32**(3), pp.378-380, March 2011. **DOI:** 10.1109/LED.2010.2099100

[28] L. Zhang, R. Huang, D. Gao, D. Wu, Y. Kuang, P. Tang, W. Ding, Y. Wang and A. Wang, "Unipolar Resistive Switch Based on Silicon Monoxide Realized by CMOS Technology", IEEE Electron Device Letters, **30**(8), pp.870-872 ,August 2009. DOI: 10.1109/LED.2009.2024650

[29] R. Ma, Q. Chen, W. Zhang, F. Lu, C. Wang, A. Wang, Y. H. Xie and H. Tang, "A Dual-Polarity Graphene NEMS Switch ESD Protection Structure", IEEE Electron Device Letters, **37**(5), pp. 674-676, May 2016. DOI: 10.1109/LED.2016.2544343

[30] W. Zhang, R. Ma, Q. Chen, M. Xia, J. Ng, A. Wang and Y-H. Xie, "The electro-mechanical responses of suspended graphene ribbons for electrostatic discharge applications", Applied Physics Letter, **108**, 153103 (2016). *http://dx.doi.org/10.1063/1.4946007*

[31] A. Wang, "Novel 3D 4-Gate MOSFET and 5-ary Logic for Hi-density Hi-speed Memory", a proposal to SRC, Sept. 2003.

[32] C. Li, F. Zhang, M. Di, Z. Pan and A. Wang, "Advances in 3D Heterogeneous Structures and Integration for Future ICs", in Proc. IEEE SOI-3D-Subthreshold Microelectronics Technology Unified Conference (S3S), 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 6, 270-279 (2019)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Multiscale Texture Analysis and Color Coherence Vector Based Feature Descriptor for Multispectral Image Retrieval

Devulapalli Sudheer, Rajakumar Krishnan*

School of Computer Science and Engineering, Vellore Institute of Technology, 632014, india

ARTICLE INFO

Article history:
Received: 07 September, 2019
Accepted: 28 November, 2019
Online: 16 December, 2019

Keywords:
CBIR
Texture
Feature Extraction
Remore Sensing
Curvelet
CCV

ABSTRACT

Content Based Image Retrieval (CBIR) for remote sensing image data is a tedious process due to high resolution and complexity of image interpretation. Development of feature extraction technique is a major portion to represent the image content in an optimal way. In this paper, we propose a feature descriptor which combines the color coherent pixel information and GLCM texture features in multi scale domain. Curvelet transform is used to decompose the image into coarse and detail coefficients. Then Gabor magnitude is computed for each coefficient to improve the directional information. GLCM texture features are extracted from the Gabor magnitude response. The novel feature set by combining the CCV and GLCM using curvelet and Gabor filter is developed. Mahalanobis distance measure is used to find the similarity between query and feature database. Average Normalized Modified Retrieval Rate (ANMRR) is computed to evaluate the performance with the state of art methods.

1 Introduction

Texture represent the pattern arrangement of the pixels in the image. It is an inherent property of any image. Especially in remote sensing applications, images represent the surface reflectance of the earth. Various sensors are used to capture the earth surface like Multi/Hyperspectral, RADAR and LiDAR [1]. Texture extraction in spatial domain using Local Binary Patterns (LBP) and Gray Level Co-Occurrence Matrix has been explained in [2, 3, 4]. Other local texture pattern features like Local Derivative Patterns (LDP), Local Tetra Patterns (LTrp) are reviewed by [5, 6]. Recent advancements in texture features have been developed using multi-scale GLCM windows [7] and Local Contourlet Tetra Patterns [8]. Texture analysis in multi-scale and multi resolution analysis is very useful to extract more precise features. In proposed work, Curvelet transform has been used to analyze the input image in multi scale domain. Curvelet is directionally more efficient than other existing multi resolution transforms. It can preserve the tiny non linear edge information due to its anisotropic nature. Coarse and detail scale coefficients of curvelet are used to compute the Gabor magnitudes. Total four Gabor magnitude responses are be used to extract the GLCM properties. Applying the Gabor magnitudes on each scale of curvelet gives more directional features of the non linear edges. Texture information extracted from these coefficients can perform retrieval applications efficiently. To carry the spatial domain information, color coherent pixels are computed and merged to the texture feature descriptor. The proposed feature descriptor outperforms the state of art retrieval methods. It has been tested on the public remote sensing image retrieval dataset namely UCmerced.
The paper has been organized as follows: Section 2 Related work, Section 3 Methodology,Section 4 Results & Evaluation and Section 5 Conclusion.

2 Related Work

2.1 Importance of Texture Feature in Remote Sensing

In recent years, many texture-based feature descriptors have been developed for the purpose of CBIR in remote sensing. This section explaines comprehensive review of texture feature extraction in remote sensing image data. The importance of spatial information retrieval in theoretical perspective for model fit and model selection techniques has been explained by Mihai Datcu, et.al, in (1998) [9]. Seisuke Fukuda, et.al (1999) have developed a wavelet-based texture feature set for the polarimetric SAR image data, Where Downsampling of

*Corresponding Author: Rajakumar Krishnan, Vellore Institute of Technology, Vellore, +91 9894079434 & rajakumar.krishnan@vit.ac.in

wavelet coefficients are omitted. The energy of the wavelet coefficients is considered as texture descriptor [10]. M. Chica-Olmo, et.al (2000) have proposed univariate and multivariate texture features by using spatial variability based variogram measures [11]. It will refine the spatial and spectral structure of the image. A multiband image will be given as an output to quantify the local spatial variability of radiometric data. Jean-Luc Starck, et.al (2002) have improved the curvilinear feature extraction in multiscale domain by using curvelet transform to overcome the wavelet drawbacks [12]. Second and higher order statistical measures have been applied to get the texture information in multiscale and multi resolution domains. Andrey V. Bogdanov, et.al (2005) have applied neural network based classification on GLCM features for the classification of the sea ice in multi-sensor data [13]. F. Kayitakire (2006), et.al, have derived Forest variables from the high spatial resolution data by using GLCM [14]. Interactive retrieval system is proposed by Marin Ferecatu, et.al (2007), using relevance feedback method to improve the accuracy [15].

Feature level fusion also improves classification accuracy. Vijaya V, et.al (2009) had developed a system for texture feature classification by fusing statistical and wavelet-based texture features [16]. Changren Zhu, et.al (2010) have proposed shape and texture based feature set for detecting the ship in spaceborne optical images. Compactness, Convexness, Rectangularity, and Eccentricity are used for the shape based feature set and Mean, Standard variance and Entropy are used for analyzing the statistical features. Local Multiple Patterns are used as local texture features [17]. Shufu Xie, et.al (2010) had developed a Local Gabor Xor Patterns (LGXP) for facial recognition [18]. Multi-resolution based feature extraction has succeeded in extracting robust features. Gholamreza Akbarizadeh, et.al (2012) have proposed Kurtosis wavelet energy feature for the texture recognition in SAR images [19]. 3D DWT is suitable for extracting structure and texture information from the Hyperspectral data. Hyperspectral image consists of continuous spectral bands that provide more accurate information. Yuntao Qian, et.al (2012) has proposed a 3D wavelet-based texture feature extraction and classified it using sparse linear regression [20]. Simple statistics, Gabor filter bank and color histogram has been used as feature descriptors for efficient image retrieval by Yi Yang, et.al (2012) [21]. Local spectral histogram features are generated by merging the local histograms. Jiangye Yuan, et.al, (2014) have computed local spectral histograms and Gabor filter based texture features to segment the image regions [22]. Morphological texture features based on structuring elements have also shown good results in CBIR techniques. Erosion, Dilation, Open, Close operations by structuring elements will improve the feature extraction efficiency. Erchan Aptoula (2014) have developed a morphological texture descriptor based on Circular Covariance Histogram (CCH) and Rotation Invariant Triplet (RIT) for Land Cover image retrieval [23]. Alaa Al-Hammami and Hisham Al-Rashdan (2010) have proposed an enhancement to the Color Coherence Vector (CCV) based on modification of the distance and angle [24]. Zhenfeng Shao, et.al have extracted Gabor filter based texture and color features for retrieval application. Unichrome feature of each channel of the image data using Gabor filter with orientation (u, v) and opponent feature with (u, v') are used as color features respectively [25]. GLCM is the best features to extract texture information of mono channel image data. For Multichannel or Hyperspectral image, Data clustering and sparse representation based GLCM has been proposed by Xin Huang, et. al (2014) [26]. Local Binary Patterns (LBP), Local Ternary Patterns (LTP), Local Terra Patterns (LTrP) and Local Derivative Patterns (LDP) are most popular texture extraction methods evaluated by [27, 28, 29]. Color and Texture descriptor based image retrieval system has been developed by using Block Difference Inverse Probability (BDIP) and Block Variation of Local Correlation Coefficients (BVLC) by Chandan Singh, et.al (2016) [30]. Spectral end members and texture features based on GLCM has been developed for hyperspectral image retrieval by Jing Zhang, et.al (2017) [31].

2.2 *Efficiency of Curvelet Transform*

The wavelet-based approach have failed to preserve nonlinear singularities due to the directional limitation. Anisotropic nature of curvelet transform helps in extracting curved singularities. It is a multiscale and multi-resolution transform which can process an image in several orientations at each scale. Curvelet-based SAR image classification has been performed by using the Histogram of Curvelet (HOC) [32]. The curvelet transform has been developed by Emmanuel J. Candes, David L. Donoho, et al. (2002) in order to overcome the drawbacks of wavelet [33]. Lindsay Semler et al. (2006) highlighted the curvelet-based texture descriptors over CT scan images. Energy, entropy, mean, standard deviation has been calculated by using the effective curvelet-based texture descriptors. The results have been compared with similar algorithms based on wavelet and ridgelet descriptors [34]. Tobias Gebacket al. (2009) have explained that when the number of coefficients that are required to represent the edge to a specified accuracy is regarded, curvelets dispense an almost ideal representation of C^2 edges. Decomposition of the image f into J feature levels have been furnished by the discrete curvelet transforms, with L_j directions on each level, and $K_{jl,1} \times K_{jl,2}$ spatial shifts for each of these directions. [35]. Jianwei ma and Gerlindplonka (2010) have described that the compression based on a wavelet, removal of noise or structure extraction have become computationally inefficient for geometric features with line and surface singularities. The first-generation curvelets have been used for image denoising. The first-generation curvelets are also been used to intensify image contrast, render astronomical images and integrate satellite images.After introduction of the second-generation curvelets in 2004, the applications of curvelets escalated in several fields. Hybrid methods are more popular that are pertaining to discrete curvelet transform, which is formulated by incorporating curvelets with another technique for image processing [36]. F. Gomez et al. (2011) have mapped the image to the curvelet space thus data diffusion is approached by

statistical parametric model for each of the sub-bands. Their proposed descriptor (Rotation invariant texture characterization using a curvelet-based descriptor) have been calculated by using the statistical patterns of the curvelet coefficients. Kullback-Leibler distance have been used to find the distance between statistical features. Curvelet is efficient in multiscale decomposition to improve the limits in wavelets for representation of geometrical information [37]. Yungang Zhang et al. (2011) had described that the idea of Discrete Curvelet Transform (DCT) is easy to choose acceptable sampling at the range of scales, position, and orientation. The prevalence space is decomposed into dyadic rectangular coronae. These dyadic spaces are further divided into wedges. The number of wedges increases two times at each level. [38]. NourElDin Laban et al. (2012) have made use of two dissimilar types of features, viz. Color Histograms (CH) and textures in wavelet domain for every band of the multispectral image. The quantization method is use to extract the color histograms from satellite images. Daubechies wavelets are used to extract texture features. The above authors proposed a technique to improve the procedure of satellite image retrieval using two stages: Candidate selection stage and Refinement stage. The refinement stage depends on the outcomes of the candidate selection stage that comprises the adjoining regions in which non-uniform polygon shapes that cover the semantic type area are utilized. They have evaluated the potentiality of their system using a wide area of satellite images and then analyzed in terms of precision and recall rates [39]. ErkanUslu et al. (2014) have stated that the curvelet becomes finer and smaller in the spatial domain and shows more reactivity to curved edges when the resolution level increases. This allows to capture the curves in an image efficiently. As a result, a few coefficients can estimate curved singularities. The energy feature is computed in both discrete curvelet transform and wavelet transform[40]. Benoit Beguet et al. (2014) have developed a global framework of forest variable retrieval methodology by using Haralick Texture Features. The panchromatic and multispectral bands are used to create test frames which are made up of image samples. Then the texture features which have several parameterizations are computed on test frames. Using these test frames, linear regression based feature selection is performed for retrieving the best texture features pertaining to each forest variable. Both single and multiple linear regressions are processed. Multi-scale texture descriptors are formed by choosing the best image texture features through statistical modeling. Five common forest structure variables have been modeled efficiently [41]. Wen-Chyi Lin et al. (2015) described that the curvelet transform for curved singularities in images is particularly attractive for multiscale analysis of microscopic tissue images where cellular structural information exhibits abundant curved singularities. Mean, energy, variance, and entropy vales are calculated as texture features in each sub-band. SVM classifier is used to classify the results of [42]. Wen-Chyi Lin et al. (2016) have explained that higher-order statistical moments of fine-scale curvelet coefficients are selected as critical features. To incorporate the skewness and kurtosis of curvelet coefficients at a specified scale as texture features. The above authors added the third and fourth order moments to variance, energy, and entropy, so that a better discriminative capability is provided to them. The curvelet coefficients are chosen as discriminative features because they are used to implicitly capture the nucleus and glandular boundary [43]. Sudheer Devulapalli and Rajakumar Krishnan (2019) have developed a image fusion application for multi sensor image data using curvelet coefficients and adaptive neuro fuzzy inference system [44]. Du-Ming Tsai et al. (2017) states that a structural texture is depicted in the satellite image by planting pattern of coffee trees. To extricate the structural features in the spectral domain for image segmentation, a technique based on Fourier transform is proposed. Unidirectional frequency components with high energy are produced by row planted coffee fields, whereas the frequency components are produced in different directions in the spectral domain image by naturally growing plants. The experimental outcome shows that row-planted coffee field can be sectioned in a satellite image with the conjunction of structural and density features. The yield of coffee production can be determined from the data on physical area and density, which is provided by these features. [45].

3 Methodology

Remote sensing benchmark multispectral dataset UC Merced is considered for the experiment in the proposed work. The dataset is consisting of 21 classes of land use data for research purpose [46]. The images are subsampled from USGS National Map urban area image collection. Each image consists of 256 x 256 pixels. The resolution of each pixel is 1 foot. The input image bands will be separated to apply the 2D Fast Discrete Curvelet Transform in Unequally Spaced Fast Fourier Transform (USFFT). Curvelet transform will decompose the image into three coefficients which are coarse, detail, and fine scale coefficients. The detail scale coefficients again consist of multiple orientations at each resolution [47]. Figure. 1 shows the process of retrieval system.

3.1 Feature Extraction

3.1.1 Probabilistic Principle Component Analysis (PPCA)

PCA is a popular method for dimensionality reduction. It will increase the variance of the projected space. The high dimensional strongly correlated data $\{D_p\}_{p\epsilon P}$ will be converted into low dimensional uncorrelated data $\{D_l\}_{l\epsilon L}$ removing low variance data. This may lead to some data loss due to reducing the dimension to (P-L) Eigen subspace. PPCA is a method of finding principal axes by the probabilistic way [48]. Each variable of this model is calibrated by a specific probability density function (pdf). The uncorrelated input variable z with some additive noise e is represented by the equation form x=A.z+e. Where A is model parameter. If some probability densities are set to the generative input variable then it can be called as probabilistic model. PPCA is used to find the maximum likelihood parameter set $\Theta \equiv A, \lambda$. Where $A\epsilon\mathbb{R}^{(p\times L)}$ is

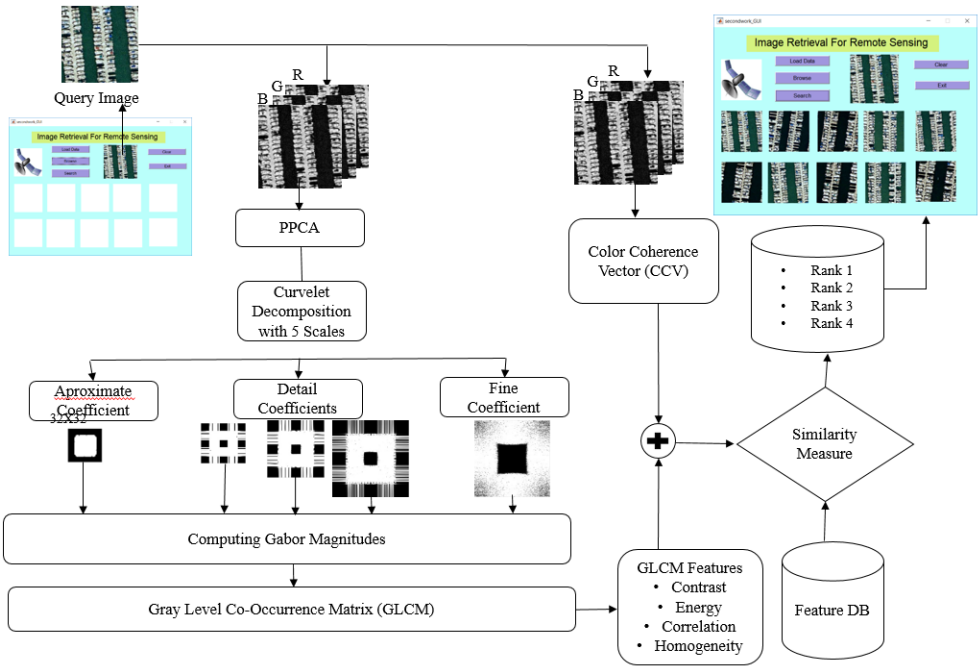


Figure 1: Architecture of the proposed system

the loading the matrix, z is pdf of latent variable and e is pdf of additive noise variable. The parameter Θ can be measured by Expectation and Maximization (EM) algorithm [49]. PPCA assumes that $p(z) = \aleph(z : 0, I)$ and $p(e) = \aleph(e : 0, \lambda.I)$. Gaussian pdf is similar to linear operations so that the linear transformation of the Gaussian will produce the new Gaussian distribution.

3.1.2 Curvelet Decomposition

The directionality and anisotropic natures are the main features of the curvelet transform which can help to extract more information from the image data. The curvelet transform decomposes the image into three components like coarse, detail and fine scale coefficients [50]. The input image I will be decomposed with the help of low pass and bandpass filters $l_0, \Delta s$ respectively. $\Delta_s = \{\Delta_1 I, \Delta_2 I, \Delta_3 I \ldots \Delta_s I\}$ where Δ_s contains the 2^{-2s} width. The Discrete Fourier Transform (DFT) will be applied to the input image to apply the filters at different scales. Mathematical derivation of the DFT is as follows

$$f(n_1, n_2) = \sum_{p=0}^{M-1} \sum_{1=0}^{N-1} f(p,q) e^{-i2\pi(\frac{n_1 p}{M}, \frac{n_2 1}{N})} \quad (1)$$

The default number of scales will be log 2(n) – 3. Where n is length of the image. Curvelet coefficients are denoted as $c\{j\}\{l\}\{k\} = \langle f, \varphi_{j,l,k} \rangle$. Where j represents scale, l represents orientation and k represents the position (k_1, k_2). The curvelet coefficients can be stored in a cell matrix $C_{j,l}$. Here scale j is from coarse to finest and orientation l is starts from top left corner and rotates in clockwise[51]. The rotation angles $\theta_{j,l}$ will be equidistant and sequence can be represented as $\theta_{j,l} := \frac{\pi l_2^{-\lceil j/2 \rceil}}{2}$ with $l = 0, 1, \ldots 4.2^{\lceil j/2 \rceil} - 1$. For the value of j=4 rotation angle $\theta_{j,l} = \pi l/8, l = 0, 1, 2, \ldots 15$ and eight translations with different grids will be considered. The grid sizes will be $1/2^{-j} \times 1/2^{-j/2}$. So the anisotropic scaling principle of the curvelet will be width = length2. Each subband is partitioned into squares of proper scales as, $\Delta_s I \rightarrow (W_Q \Delta_s I)_{Q \epsilon Q_s}$. Where W_Q is smooth window. Assume that window WQ (n1, n2) is supported within the sheared rectangle $p_j = \left\{(n_1, n_2) : 0 \leq n_1 - n_0 < L_(j), -L_j/2 \leq n_2 < L_j/2\right\}$ The localized dyadic squares $Q = \left[\frac{k_1}{2^s}, \frac{(k_1+1)}{2^s}\right] \times \left[\frac{k_2}{2^s}, \frac{(k_2+1)}{2^s}\right]$. By multiplying the function with corresponding window function will give the localized result near Q. Let $(T_Q I)$ denotes the operator which can transport and renormalizes I so the part of the input supported near Q will become the part of the output supported near to $[0,1]^2$. $T_Q I(n_1, n_2) = 2^s I(2^s n_1 - k_1, 2^s n_2 - k_2)$.The outcomes of dyadic squares are renormalized into single scale. $g_Q = (T_Q)^(- 1)(W_Q \Delta_s I), Q \epsilon Q_s$. Each standardized dyadic square will be analyzed by ridgelet transform. It has basis element ρ_λ that make an orthonormal basis for $l^2(R^2).\alpha_\mu = (g_Q, \rho_\lambda) where \mu = (Q, \lambda)$.

3.1.3 Color Coherence Vector

Color histograms are the most popular features to extract color descriptors for CBIR. But the spatial information will not cover by the color histogram features. To consider the color and spatial information of the image Color Coherence Vector (CCV) [53, 52] descriptor has been used. The first stage to compute the CCV is blurring the image with Gaussian filter. Then the discretization of the color space will be applied to divide into n distinct colors of the image. The input image has

3 channels, discretization function will find unique colors in each channel to form V unique values of each channel. Table. **??** shows the pseudo code for computing CCV feature vector.

3.1.4 Gabor Magnitude

Magnitude response of image will be calculated by applying Gabor transform for an input image in multiple scales and different orientations [54]. Bandlimited filters can be used to represent optimal localization in both spatial and frequency domain. If the amplitude of the frequency spectrum goes to zero which exceeds the threshold frequency value. The Gabor transform can also be defined as a Short Time Fourier Transform (STFT) with a Gaussian kernel. The 2D Gabor filter can be mathematically defined as

$$\psi_{x,y}(u,v) = \frac{f_x^2}{\pi\gamma\eta} e^{\left(\frac{f^2}{\gamma^2}u'^2 + \frac{f^2}{\eta^2}v'^2\right)} \tag{2}$$

where $u' = ucos\theta_y + vsin\theta_y$, $v' = -ucos\theta_y + vsin\theta_y$, $f_x = \frac{f_{max}}{2^{(x/2)}}$, $\theta_y = y\pi/8$. f_x and θ_y Parameters are the center frequency and orientation of the plane wave for the Gaussian kernel. The variables γ is the ratio of the center frequency and η is the sharpness constant. If the γ and η values are fixed then the center frequency f_u will define the scale of the Gabor filter. Gabor filters are complex type, which is a combination of sine and cosine functions. Gabor filter is applied to the input gray image with wavelength 4 and 90-degree orientation. Wavelength denotes the sinusoidal carrier in pixels per cycle. Orientation defines the direction of the plane wave range in 0 to 360 degree. Convolution of the Gabor filter and input image will produce the resultant image in the time-frequency domain. It can be decomposed into real and imaginary parts [55]. The mathematical representations of the convolution and decomposition as follows

$$T_{x,y}(u,v) = I(u,v) * \psi_{x,y}(u,v). \tag{3}$$

$$E_{x,y}(u,v) = Re\Big[T_{x,y}(u,v)\Big] \tag{4}$$

$$O_{x,y}(u,v) = Im\Big[T_{x,y}(u,v)\Big] \tag{5}$$

where $Re[T_{x,y}(u,v)] = cos(2\pi(x_0u + y_0v) + P)$ and $Im[T_{x,y}(u,v)] = sin(2\pi(x_0u + y_0v) + P)$. $T_{x,y}(u,v)$ is the transformed image obtained after convoluting with Gabor filter $\psi_{x,y}(u,v)$. u, v are the spatial coordinates of the image. Where I is the input image. $E_{x,y}(u,v)$ and $O_{x,y}(u,v)$ are the even and odd functions for the sin and cosine functions and P is the phase response of the sinusoid carrier. From the decomposed sine and cosine functions of the Gabor filter, Magnitude and Phase responses will be calculated as follows

$$M_(x,y)(u,v) = \sqrt{\left(E_{x,y}^2(u,v) + O_{x,y}^2(u,v)\right)} \tag{6}$$

$$\phi_{x,y}(u,v) = ArcTan(O_{x,y}^2(u,v)/E_{x,y}^2(u,v)) \tag{7}$$

The magnitude responses for the curvelet coefficients are considered for the extraction of Haralick texture features. The adjoint transform of the curvelet coefficients will be computed for each scale j and orientation l by applying 2D Inverse Fast Fourier Transform (IFFT).

$$g_{j,l}[n_1,n_2] = F(C_{j,l,k}) \tag{8}$$

$$g[t_1,t_2] = IFFT\left(\sum_{j=1}^{J}\sum_{l=0}^{L} g_{j,l}[n_1,n_2] * U_j[n_1,n_2]\right) \tag{9}$$

The adjoint transform will be computed for coarse, detail and fine scale coefficients separately. Gabor transform will be applied on the resultant image at each scale to compute the magnitude response. The surface view of the magnitude response at each scale has shown in Figure. 2.

3.1.5 Texture Analysis

To identify local or global pixel patterns, texture features are very important. GLCM plays a vital role in texture analysis. Haralick, et.al, (1973) derived statistical texture features based on GLCM. GLCM matrix consists value i with a sum of the frequency of the pixel value at a pixel value u with a pixel value y in a specific spatial orientation at v. Mathematically it can be represented as follows

$$G_{u,v} = Freq(F(x,y) = u \& F(x+1,y+1) = v) \tag{10}$$

The texture is an inherent property for all kind of surfaces. It can give important information about the pixel arrangement structure of each surface and spatial relationship. The smallest differences in texture for large image databases is very difficult. GLCM properties will be helpful to detect the texture features in local and global areas. Some of the GLCM properties derived by Haralick [4] are considered for the experiment.

$$Energy = \sum_{u=0}^{n-1}\sum_{v=0}^{n-1} GLCM(u,v)^2 \tag{11}$$

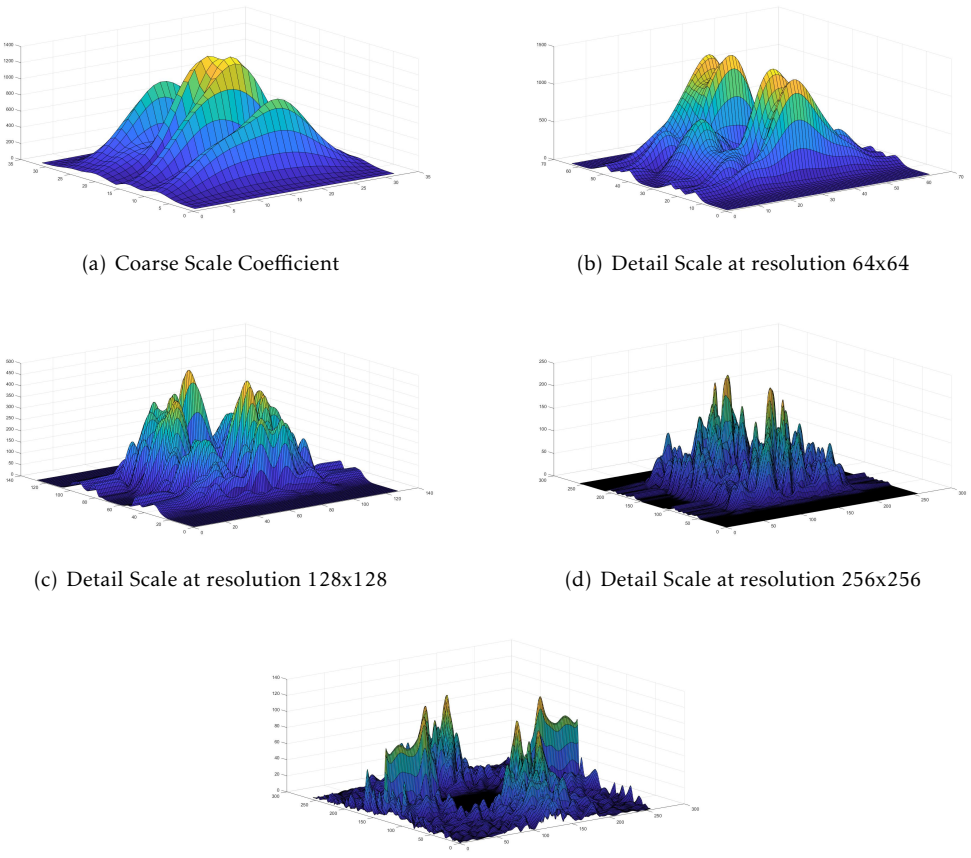
$$Correlation = \sum_{u=0}^{n-1}\sum_{v=0}^{n-1} \frac{GLCM(u,v)(u-\mu_x)(v-\mu_y)}{\sigma_x\sigma_y} \tag{12}$$

$$Contrast = \sum_{u=0}^{n-1}\sum_{v=0}^{n-1} GLCM(u,v)(u-v)^2 \tag{13}$$

$$Homogeneity = \sum_{u=0}^{n-1}\sum_{v=0}^{n-1} \frac{GLCM(u,v)}{1+(u-v)^2} \tag{14}$$

3.2 Similarity Measure

In this work, the Mahalanobis distance measure is used as a distance metric. Suppose if there are two groups G_1 and G_2 with some characteristics. these characteristics are considered as the feature vectors representing the objects of each group



(a) Coarse Scale Coefficient

(b) Detail Scale at resolution 64x64

(c) Detail Scale at resolution 128x128

(d) Detail Scale at resolution 256x256

(e) Fine Scale at resolution 256x256

Figure 2: Gabor Magnitude Responses of the Curvelet Coefficients

[56]. the feature vector x with dimension p have the same variation about its mean within either group. The mathematical representation of Mahalanobis distance is as follows

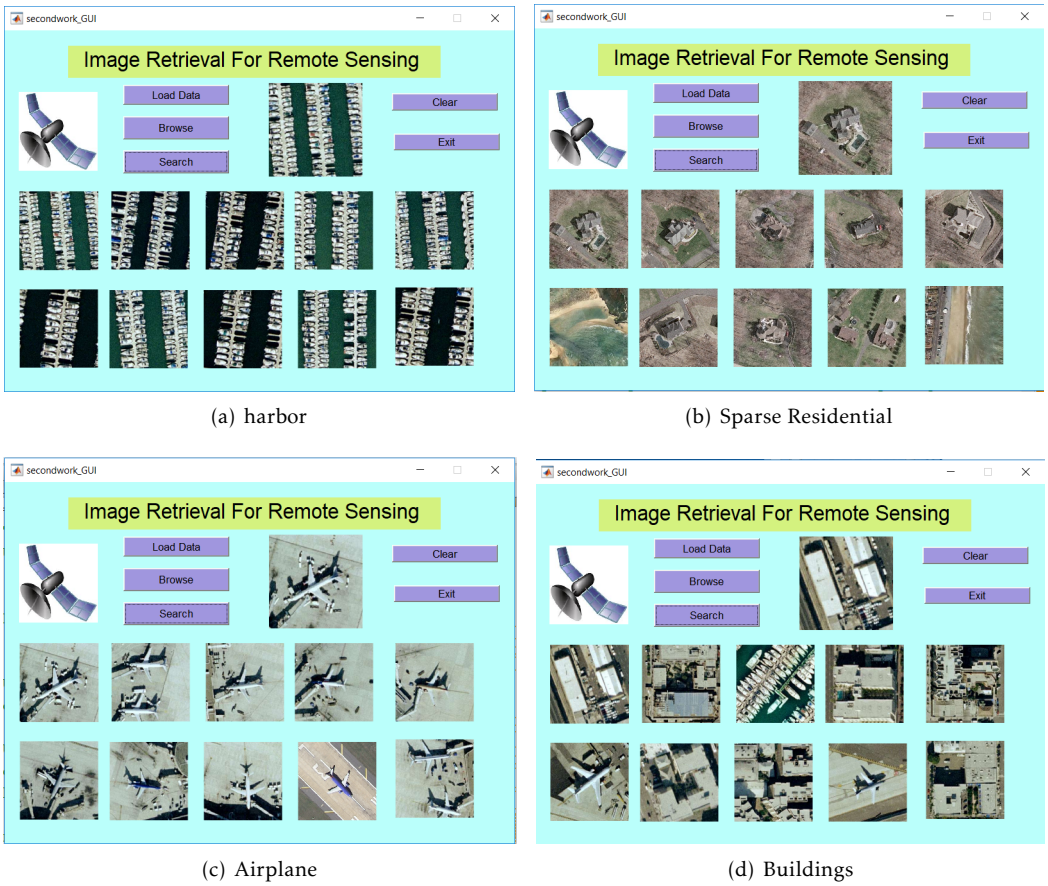
$$Dist = (\mu_1 - \mu_2)^T S^{-1} (\mu_1 - \mu_2) \tag{15}$$

where $\mu_1 and \mu_2$ are the mean of two vectors S is covariance matrix.

4 Results and Evaluation

The features have been extracted using the process explained in the section2. The experiment has been conducted on UC Merced Land use dataset which contains 21 classes with 50 samples of each class. Each image has 256×256 pixels. The spatial resolution of the data is 0.3m. The feature descriptor size for each image is 1×24. color coherence vector feature length is 1×8 because in this experiment only coherent pixel information is considered as a feature. The color coherence vector feature is computed for the multispectral image directly. Further feature extraction process the input image will be reduced to a single band by applying PPCA. The resulting of PPCA image will be decomposed into coarse and detail coefficients using curvelet transform. The coarse coefficient in one scale and detail coefficients at three different scales are used to extract the texture information. Obtained Gabor magnitude response of each coefficient to compute the GLCM properties. Contrast, Correlation, Energy and Homogeneity features have been extracted from each coefficient. The feature length of texture features is 1×16. The color and texture features will be combined together to form a single feature descriptor of an image. Feature descriptor of each image in the database will be computed and stored as a feature base. While the query image submitted to the retrieval system the same features will be extracted to describe the content of the image and matched with the feature base using Mahalanobis similarity metric. The feature obtained lower distance is considered as the most relevant image to the query. In this experiment, top 10 relevant images are shown as a result. Figure. 3 shows the retrieval results of four different classes in the dataset.

The performance metrics are very much important to compare the system accuracy with other systems. The most important performance measures are precision and Recall. Precision can be described as a number of relevant images retrieved over a total number of images retrieved. Recall can be defined as number of relevant images retrieved over the total number of relevant images in the database. These traditional mea-



(a) harbor (b) Sparse Residential

(c) Airplane (d) Buildings

Figure 3: Retrieval Results of the System

sures will not considered the position of the relevant image appears among the top retrieved results. The retrieval system performance is evaluated by using Average Normalized Modified Retrieval Rate (ANMRR) metric [57, 58]. ANMRR The performance metric ANMRR as shown below

$$R(I) = \begin{cases} R(I) & if\ R(I) \le I(q) \\ 1.25 * I(q) & if\ R(I) > I(q) \end{cases} \quad (16)$$

where R(I) is position of relevant image I retrieved, I(q) is retrieved images for query q. $I(q) = min\{4.TG(q), 2.max[TG(q), \forall q]\}$ Average Rank will be calculated as follows

$$Average_R(q) = \frac{1}{TG(q)} \sum_{I=1}^{TG(q)} R(I) \quad (17)$$

where TG(q) is the total ground truth images in the database to the query q. TG(q) is influencing the average rank in the above equation. to reduce this factor Modified Retrieval rate (MRR) is computed as $Aerage_R(q) - 0.5.[1 + TG(q)]$. Normalization will be applied to make the value exist between 0 and 1.

$$NMRR(q) = \frac{MRR(q)}{1.25 * I(q) - 0.5 * [1 + TG(q)]} \quad (18)$$

Finally the ANMRR is derived for the Total number of queries TQ is follows

$$ANMRR(q) = \frac{1}{TQ} \sum_{q=1}^{NQ} NMRR(q) \quad (19)$$

Table 2 shows the ANMRR values for the sample retrieval results with ten queries from each class. we have tested for the four classes in the database. Retrieval performance was compared with the five existing popular texture feature extraction methods. The proposed method is developed by combining the CCV and texture features using GLCM from the Gabor magnitude of the curvelet sub-bands in different scales. Extracting Gabor magnitude from the curvelet sub-bands in multiple scales will improve the directional features of the input image. Curvelet is capable of extracting anisotropic edge information in various orientations at different scales. The resultant of the curvelet decomposed image will have one coarse and three detail coefficients as explained in the section 3. The comparison of the proposed method with existing literature has been shown in Figure 4.

5 Conclusion

This paper develops a feature extraction method which is a combination of color coherent pixel information and multi

	Harbor	Sparse Residential	Airplane	Buildings
LBP	0.2405	0.2633	0.2367	0.2304
LDP	0.2329	0.2633	0.2418	0.2304
LSP	0.2291	0.2658	0.245	0.2430
LTrp	0.2456	0.2646	0.2278	0.2253
Gabor [25]	0.2370	0.2642	0.2378	0.2322
CCV [24]	0.231	0.262	0.231	0.243
GLCM [26]	0.241	0.236	0.223	0.234
Proposed	0.2096	0.215	0.2096	0.2215

Table 2: ANMRR Performance Results

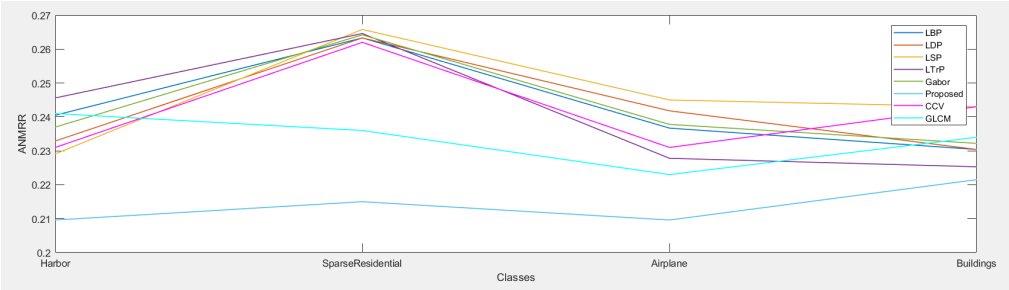


Figure 4: Performance Graph for state of art methods

scale texture features. The proposed feature descriptor is extracted using Haralick texture features from curvelet coefficients. To extract the robust directional infornamtion, Gabor magnitude responses are computed from the curvelet coefficients. Haralick texture features extracted from the resultant gabor magnitude responses are shown best retrieval rate than other state of art feature extraction techniques. It improves the retrieval accuracy of the CBIR search engines. Retrieval system was tested on various classes of remote sensing image data. ANMRR performance metric gives the best average retrieval rate for a set of queries from various classes. Four classes have been tested in the dataset. This work will be enhance by adopting machine learning techniques for the target classification for high resolution imagery data.

References

[1] Melesse, Assefa, et al. "Remote sensing sensors and applications in environmental resources mapping and modelling." Sensors 7.12 (2007): 3209-3241.

[2] He, Dong-Chen, and Li Wang. "Texture unit, texture spectrum, and texture analysis." IEEE transactions on Geoscience and Remote Sensing 28.4 (1990): 509-512. DOI:10.1109/TGRS.1990.572934.

[3] Guo, Zhenhua, Lei Zhang, and David Zhang. "A completed modeling of local binary pattern operator for texture classification." IEEE transactions on image processing 19.6 (2010): 1657-1663. DOI: 10.1109/TIP.2010.2044957.

[4] Haralick, Robert M., and Karthikeyan Shanmugam. "Textural features for image classification." IEEE Transactions on systems, man, and cybernetics 6 (1973): 610-621.

[5] Juneja, Komal, et al. "A survey on recent image indexing and retrieval techniques for low-level feature extraction in CBIR systems." 2015 IEEE International Conference on Computational Intelligence & Communication Technology. IEEE, 2015. DOI:10.1109/CICT.2015.92

[6] Yadav, Susheel Kumar, Dhirendra Pratap Singh, and Jaytrilok Choudhary. "A Survey: Comparative Analysis of Different Variants of Local Binary Pattern." 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT). IEEE, 2018.DOI:10.1109/ICICCT.2018.8473220.

[7] Lan, Zeying, and Yang Liu. "Study on multi-scale window determination for GLCM texture description in high-resolution remote sensing image geo-analysis supported by GIS and domain knowledge." ISPRS International Journal of Geo-Information 7.5 (2018): 175.DOI:10.3390/ijgi7050175.

[8] Kumar, TG Subash, and V. Nagarajan. "Local contourlet tetra pattern for image retrieval." Signal, Image and Video Processing 12.3 (2018): 591-598.DOI:10.1007/s11760-017-1197-1.

[9] Datcu, Mihai, Klaus Seidel, and Marc Walessa. "Spatial information retrieval from remote-sensing images. I. Information theoretical perspective." IEEE transactions on geoscience and remote sensing 36.5 (1998): 1431-1445.DOI:10.1109/36.718847.

[10] Fukuda, Seisuke, and Haruto Hirosawa. "A wavelet-based texture feature set applied to classification of multifrequency polarimetric SAR images." IEEE Transactions on Geoscience and Remote Sensing 37.5 (1999): 2282-2286.DOI:10.1109/36.789624.

[11] Chica-Olmo, M., and F. Abarca-Hernandez. "Computing geostatistical image texture for remotely sensed data classification." Computers & Geosciences 26.4 (2000): 373-383. DOI:10.1016/S0098-3004(99)00118-1.

[12] Starck, Jean-Luc, Emmanuel J. Candès, and David L. Donoho. "The curvelet transform for image denoising." IEEE Transactions on image processing 11.6 (2002): 670-684. DOI: 10.1109/TIP.2002.1014998.

[13] Bogdanov, Andrey V., et al. "Multisensor approach to automated classification of sea ice image data." IEEE Transactions on geoscience and remote sensing 43.7 (2005): 1648-1664.DOI: 10.1201/9781420003130.ch25.

[14] Kayitakire, François, C. Hamel, and Pierre Defourny. "Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery." Remote sensing of environment 102.3-4 (2006): 390-401. DOI:10.1016/j.rse.2006.02.022.

[15] Ferecatu, Marin, and Nozha Boujemaa. "Interactive remote-sensing image retrieval using active relevance feedback." IEEE Transactions on Geoscience and Remote Sensing 45.4 (2007): 818-826. DOI:10.1109/TGRS.2007.892007.

[16] Chamundeeswari, Vijaya V., Dharmendra Singh, and Kuldip Singh. "An analysis of texture measures in PCA-based unsupervised classification of SAR images." IEEE Geoscience and Remote Sensing Letters 6.2 (2009): 214-218.DOI:10.1109/LGRS.2008.2009954.

[17] Zhu, Changren, et al. "A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features." IEEE Transactions on geoscience and remote sensing 48.9 (2010): 3446-3456.DOI:10.1109/TGRS.2010.2046330.

[18] Xie, Shufu, et al. "Fusing local patterns of gabor magnitude and phase for face recognition." IEEE transactions on image processing 19.5 (2010): 1349-1361. DOI:10.1109/tip.2010.2041397.

[19] Akbarizadeh, Gholamreza. "A new statistical-based kurtosis wavelet energy feature for texture recognition of SAR images." IEEE Transactions on Geoscience and Remote Sensing 50.11 (2012): 4358-4368.DOI:10.1109/TGRS.2012.2194787.

[20] Qian, Yuntao, Minchao Ye, and Jun Zhou. "Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features." IEEE Transactions on Geoscience and Remote Sensing 51.4 (2012): 2276-2291. DOI:10.1109/TGRS.2012.2209657.

[21] Yang, Yi, and Shawn Newsam. "Geographic image retrieval using local invariant features." IEEE Transactions on Geoscience and Remote Sensing 51.2 (2012): 818-832.DOI:10.1109/TGRS.2012.2205158.

[22] Yuan, Jiangye, DeLiang Wang, and Rongxing Li. "Remote sensing image segmentation by combining spectral and texture features." IEEE Transactions on geoscience and remote sensing 52.1 (2013): 16-24. DOI:10.1109/TGRS.2012.2234755.

[23] Aptoula, Erchan. "Remote sensing image retrieval with global morphological texture descriptors." IEEE transactions on geoscience and remote sensing 52.5 (2013): 3023-3034. DOI:10.1109/TGRS.2013.2268736.

[24] Al-Hamami, Alaa, and Hisham Al-Rashdan. "Improving the Effectiveness of the Color Coherence Vector." Int. Arab J. Inf. Technol. 7.3 (2010): 324-332.

[25] Shao, Z., Zhou, W., Zhang, L., & Hou, J. "Improved color texture descriptors for remote sensing image retrieval." Journal of applied remote sensing 8.1 (2014): 083584. DOI:10.1117/1.

[26] Huang, Xin, Xiaobo Liu, and Liangpei Zhang. "A multichannel gray level co-occurrence matrix for multi/hyperspectral image texture representation." remote sensing 6.9 (2014): 8424-8445. DOI:10.3390/rs6098424.

[27] Fadaei, Sadegh, Rassoul Amirfattahi, and Mohammad Reza Ahmadzadeh. "Local derivative radial patterns: A new texture descriptor for content-based image retrieval." Signal Processing 137 (2017): 274-286.DOI: 10.1016/j.sigpro.2017.02.013.

[28] Tiwari, Ashwani Kumar, Vivek Kanhangad, and Ram Bilas Pachori. "Histogram refinement for texture descriptor based image retrieval." Signal Processing: Image Communication 53 (2017): 73-85. DOI:10.1016/j.image.2017.01.010.

[29] Jenitta, A., and R. Samson Ravindran. "Content Based Geographic Image Retrieval using Local Vector Pattern." Brazilian Archives of Biology and Technology 61 (2018). DOI:10.1590/1678-4324-2016160717.

[30] Singh, Chandan, and Kanwal Preet Kaur. "A fast and efficient image retrieval system based on color and texture features." Journal of Visual Communication and Image Representation 41 (2016): 225-238. DOI:10.1016/j.jvcir.2016.10.002.

[31] Zhang, Jing, et al. "Hyperspectral remote sensing image retrieval system using spectral and texture features." Applied optics 56.16 (2017): 4785-4796. DOI:10.1364/AO.56.004785.

[32] Uslu, Erkan, and Songul Albayrak. "Curvelet-based synthetic aperture radar image classification." IEEE Geoscience and Remote Sensing Letters 11.6 (2013):1071-1075. DOI:10.1109/LGRS.2013.2286089.

[33] Candès, Emmanuel J., and David L. Donoho. "New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities." Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences 57.2 (2004): 219-266. DOI:10.1002/cpa.10116.

[34] Semler, Lindsay, and Lucia Dettori. "Curvelet-based texture classification of tissues in computed tomography." 2006 International Conference on Image Processing. IEEE, 2006. DOI:10.1109/ICIP.2006.312873.

[35] Gebäck, Tobias, and Petros Koumoutsakos. "Edge detection in microscopy images using curvelets." BMC bioinformatics 10.1 (2009): 75. DOI:10.1186/1471-2105-10-75.

[36] Ma, Jianwei, and Gerlind Plonka. "A review of curvelets and recent applications." IEEE Signal Processing Magazine 27.2 (2010): 118-133.

[37] Gómez, F., and E. Romero. "Rotation invariant texture characterization using a curvelet based descriptor." Pattern Recognition Letters 32.16 (2011): 2178-2186. DOI:10.1016/j.patrec.2011.09.029.

[38] Zhang, Yungang, Wei Gao, and Jun Liu. "Integrating Color Vector Quantization and Curvelet Transform for Image Retrieval." International Journal of Design, Analysis and Tools for Integrated Circuits and Systems 2.2 (2011): 99.

[39] Laban, NourElDin, et al. "System refinement for content based satellite image retrieval." The Egyptian Journal of Remote Sensing and Space Science 15.1 (2012): 91-97. DOI:10.1016/j.ejrs.2012.05.001.

[40] Uslu, Erkan, and Songul Albayrak. "Curvelet-based synthetic aperture radar image classification." IEEE Geoscience and Remote Sensing Letters 11.6 (2013): 1071-1075. DOI:10.1109/LGRS.2013.2286089.

[41] Beguet, Benoit, et al. "Automated retrieval of forest structure variables based on multi-scale texture analysis of VHR satellite imagery." ISPRS journal of photogrammetry and remote sensing 96 (2014): 164-178. DOI:10.1016/j.isprsjprs.2014.07.008.

[42] Lin, Wen-Chyi, et al. "Curvelet-based classification of prostate cancer histological images of critical gleason scores." 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI). IEEE, 2015. DOI:10.1109/ISBI.2015.7164044.

[43] Lin, Wen-Chyi, et al. "Curvelet-based texture classification of critical Gleason patterns of prostate histological images." 2016 IEEE 6th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS). IEEE, 2016. DOI:10.1109/ICCABS.2016.7802768.

[44] Sudheer Devulapalli, Rajakumar Krishnan, "Synthesized pansharpening using curvelet transform and adaptive neuro-fuzzy inference system," J. Appl. Remote Sens. 13(3), 034519 (2019), doi: 10.1117/1.JRS.13.034519.

[45] Tsai, Du-Ming, and Wan-Ling Chen. "Coffee plantation area recognition in satellite images using Fourier transform." Computers and electronics in agriculture 135 (2017): 115-127. DOI:10.1016/j.compag.2016.12.020.

[46] Yang, Yi, and Shawn Newsam. "Bag-of-visual-words and spatial extensions for land-use classification." Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems. ACM, 2010. DOI:10.1145/1869790.1869829.

[47] Luo, Jing, et al. "Fingerprint classification combining curvelet transform and gray-level cooccurrence matrix." Mathematical Problems in Engineering 2014 (2014). DOI:10.1155/2014/592928.

[48] Tipping, Michael E., and Christopher M. Bishop. "Probabilistic principal component analysis." Journal of the Royal Statistical Society: Series B (Statistical Methodology) 61.3 (1999): 611-622. DOI:10.1111/1467-9868.00196.

[49] Ju, Fujiao, et al. "Image outlier detection and feature extraction via L1-norm-based 2D probabilistic PCA." IEEE Transactions on Image Processing 24.12 (2015): 4834-4846. DOI:10.1109/TIP.2015.2469136.

[50] Ma, Jianwei, and Gerlind Plonka. "The curvelet transform." IEEE signal processing magazine 27.2 (2010): 118-133. DOI:10.1109/MSP.2009.935453.

[51] Donoho, David L., and Mark R. Duncan. "Digital curvelet transform: strategy, implementation, and experiments." Wavelet applications VII. Vol. 4056. International Society for Optics and Photonics, 2000. DOI:10.1117/12.381679.

[52] Roy, Kalyan, and Joydeep Mukherjee. "Image similarity measure using color histogram, color coherence vector, and sobel method." International Journal of Science and Research (IJSR) 2.1 (2013): 538-543.

[53] Pass, Greg, Ramin Zabih, and Justin Miller. "Comparing Images Using Color Coherence Vectors." ACM multimedia. Vol. 96. 1996. DOI:10.1.1.29.9596.

[54] Lee, Jen-Chun. "Copy-move image forgery detection based on Gabor magnitude." Journal of Visual Communication and Image Representation 31 (2015): 320-334.

[55] Struc, Vitomir, and Nikola Pavesic. "From Gabor magnitude to gabor phase features: tackling the problem of face recognition under severe illumination changes." Face Recognition. IntechOpen, 2010.

[56] De Maesschalck, Roy, Delphine Jouan-Rimbaud, and Désiré L. Massart. "The mahalanobis distance." Chemometrics and intelligent laboratory systems 50.1 (2000): 1-18. DOI:10.1016/S0169-7439(99)00047-7.

[57] Zier, Detlef, and Jens-Rainer Ohm. "Common datasets and queries in MPEG-7 color core experiments." ISO/IEC JTC1/SC29/WG11 MPEG99 M 5060 (1999).

[58] Cieplinski, Leszek. "MPEG-7 color descriptors and their applications." International Conference on Computer Analysis of Images and Patterns. Springer, Berlin, Heidelberg, 2001. DOI:10.1007/3-540-44692-3_3.

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 272-279 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Optimization of Low Temperature Differential Stirling Engine Regenerator Design

Hind El Hassani[*,1], Nour- Eddine Boutammachte[2], Sanae El Hassani[3]

[1]ENSA Engineering School, Industrial Department, Sidi Mohamed Ben Abdellah University, 30050, Morocco

[2]ENSAM Engineering School, Energy Departement, Moulay Ismail University, 50000, Morocco

[3]ENSA Engineering School, Chouaib Doukkali University, 24000, Morocco

ARTICLE INFO

Article history:
Received: 06 January, 2020
Accepted: 10 March, 2020
Online: 20 March, 2020

Keywords:
Stirling engine
Regenerator
Optimization
Pressure drop
Heat transfer coefficient

ABSTRACT

Stirling engines working with relatively low temperature are attractive for the future, especially for applications like water pumping without concentration system and low temperature heat recovery. The present work is the continuation of a series of works on the optimization of this type of machines. The main objective of this paper is to investigate the performance of low temperature differential Stirling engine (LT-SE) regenerator by adopting a theoretical model based on heat transfer and frictional pressure drop correlations. Since the speed and the Reynolds number of this type of machines are low, the correlations concerning pressure drop were validated by comparing the calculated results with experimental measurements on a LT-SE prototype. Based on this model, several calculations were conducted in order to know how LT-SE designers can attain a desired value for regenerator effectiveness. In fact, the effect of six parameters on regenerator performances was investigated. Studied parameters are: engine speed, regenerator volume, regenerator porosity, wires diameter, working fluid type and regenerator fibers arrangement. The effect of these parameters was especially checked on pressure drop, heat transfer coefficient, regenerator and engine efficiencies. Results indicated many designing recommendations for LT-SE having the same power range of our prototype: a regenerator length around L_{reg}= 1.41 × displacer stroke, a wire diameter around d_w= 0.15mm, a porosity around β= 0.75. Results indicated also that the most performant working fluid is Helium.

1. Introduction

The Stirling engine is one of the methods for converting heat into mechanical energy with a theoretically maximum efficiency [1,2]. This is due to the use of the regenerator, which is a heat exchanger acting as an energy economizer [3]; It is a porous matrix, with a great storage capacity of thermal energy and a great heat transfer surface [4]. Stirling engine regenerators are generally composed either from woven screens or random fibers [5].

During half of the Stirling motor cycle, heat is transferred from the working fluid to the matrix. During the other half, the heat is transferred in the opposite direction. Thus, for a whole cycle, the transferred heat between working fluid and the matrix is null.

The very first mathematical theories for regenerator's functioning description were published around 1920, more than 100 years after Stirling engines invention. Later, experimental and theoretical studies of regenerators were conducted in order to identify thermodynamic phenomena acting in this important component, and to predict Stirling engines behavior. Examples of these studies include those of Kays and London [6], who presented correlations between parameters based on experimental studies in order to find out the working fluid Darcy friction factor. Later, Sodré and Parise [7] conducted experimental tests in order to identify the pressure drop in an annulus duct filled of a woven matrix. Similarly, Tanaka et al. [8] have studied flow and heat transfer characteristics of the regenerator for a periodic flow. Many numerical studies were also conducted in order to analyze the flow through regenerator matrices based on 3D CFD Modeling [9]. The finite element method seems suitable for such analysis as indicated by Rühlich and Quack [10], Gedeon and Wood [11], Ibrahim et al.[12, 13], Tew et al. [14] and others. These studies have stressed the importance of flow simulation for understanding

*Corresponding Author: Hind El Hassani, hind_elhassani@yahoo.fr

characteristics of fluid friction. Combined experimental and numerical studies have quantified the pressure drop and heat transfer rate [15]. In addition, effort has been invested to study the design and structural parameters of mesh type regenerators of Stirling cryocoolers [16].

The present study contributes to the optimization of a category of solar low temperature differential Stirling engines (LT-SE). Prototypes were designed and constructed within a cooperation framework between Dresden University- Germany, and Moulay Ismail University- Morocco. These prototypes are designed to be simple and cheap for providing a solar pumping water system, and thus, contributing to human development [17].

Since this category of Stirling motors is to be developed and optimized, the authors studied earlier many parameters to find out there optimal values [18]. This paper aims to complete this development process by studing and optimizing one of the critical components of this type of engines: the regenerator.

In the literature, studies about regenerators are focused on high and moderate temperature differential Stirling engines. This paper aims to fill out the existing lack about regenerators of low temperature differential engines [19,15]. In this work a theoretical study of regenerators is held, taking into account the pressure drop inside it, heat transfer coefficient, regenerator and engine efficiencies. Since the speed and the Reynolds number of this type of machines are low, the correlations concerning pressure drop were validated by comparing the calculated results with experimental measurements on a LT-SE prototype.

After validation of the theoretical model, a parametrical study is conducted in order to find optimal values of a number of design and functioning parameters. The considered prototype is a beta type of LT-SE, called SUNWATER 3. Figure 1 and Table 1 present its main components and geometrical specifications.

Table 1: Technical specifications and calculation data of SUNWATER 3

Parameter	Value
Displacer diameter	1.35 *m*
Absorber diameter	1.51 *m*
Working piston diameter	0.29 *m*
Regenerator volume	0.063 m^3
Fibers arrangement	Random
Hot temperature	70°*C*
Cold temperature	20 °*C*
Working fluid	Air
Cooling system	Water cooled
Section area of the regenerator housing	≈ 0.36 m^2
Average pressure	≈1.1 *bars*

2. Theoretical analysis

In order to identify the effect of a number of regenerator parameters on its efficiency and on LT-SE efficiency, some correlations and equations were adopted, which are suitable for Stirling engines characteristics'.

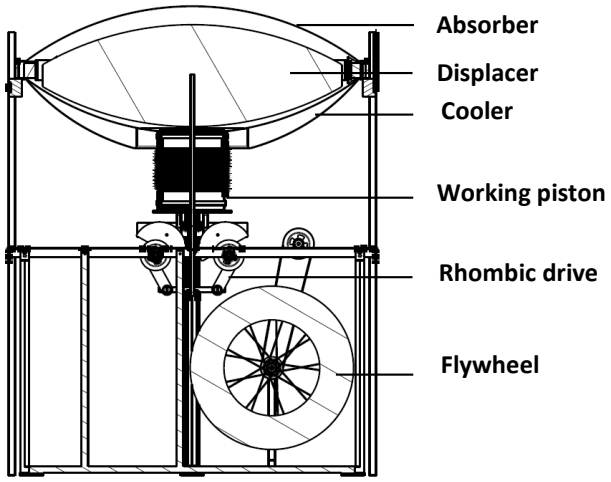


Figure 1: Illustration of SUNWATER 3

2.1 Correlations for Darcy friction factor and Nusselt number

Experimental studies of Gedeon and Wood [11] about regenerators have led to the correlations of Darcy friction factor and Nusselt number for two types of regenerators: felts and woven screens, given in equations (1 - 4):

- Felts fibers [11] :

$$f = \frac{192}{Re} + 4.53\ Re^{-0.067} \quad (1)$$

$$Nu = (1 + 1.16\ Pe^{0.66})\beta^{2.61} \quad (2)$$

- Woven screens [10]:

$$f = \frac{129}{Re} + 2.91\ Re^{-0.103} \quad (3)$$

$$Nu = (1 + 0.99\ Pe^{0.66})\beta^{1.79} \quad (4)$$

With:

- Re is the Reynolds number, defined as :

$$Re = \frac{d_h\ u_R\ \rho}{\mu}$$

- β is porosity, calculated as following :

$$\beta = \frac{\text{Volume du vide}}{\text{Volume total}}$$

- u_R is fluid velocity [20] :

$$u_R = \frac{Q_m}{\beta\ A\ \rho}$$

- d_h is the hydraulic diameter, for porous matrices, defined in [21] as:

$$d_h = \frac{\beta\ d_w}{1 - \beta}$$

- Pe and Pr are respectively Peclet and Prandtl numbers, given as :

$$\mathrm{Pe} = \mathrm{Re}.\,\mathrm{Pr}$$
$$Pr = \frac{\mu\, C_p}{k}$$

2.2. Convective heat transfer coefficient and pressure drop:

The convective heat transfer coefficient is defined based on Nusselt number as follows:

$$\mathrm{h} = \frac{\mathrm{k\,Nu}}{\mathrm{d_h}}$$

The pressure drop in the regenerator is defined by [20]:

$$\Delta \mathrm{P} = \mathrm{f} \mathrm{L_{reg}}\, \rho\, \mathrm{u_R^2}/2\mathrm{d_h} \quad (5)$$

2.3. Regenerator efficiency

Based on Tanaka et al. work [8], the regenerator efficiency can be expressed based on the number of transfer units as follows:

$$\varepsilon = \frac{\mathrm{NTU}}{2 + \mathrm{NTU}} \quad (6)$$

with:

$$\mathrm{NTU} = \frac{\bar{\mathrm{h}}\, \mathrm{A_{wg}}}{\mathrm{C_p} \bar{\mathrm{Q}}_\mathrm{m}} \quad \text{and} \quad \mathrm{A_{wg}} = \frac{4\,(1-\beta)\mathrm{V}}{\mathrm{d_w}}$$

$\bar{\mathrm{h}}$ et $\bar{\mathrm{Q}}_\mathrm{m}$ are average values of convective heat transfer coefficient and mass flow rate respectively.

2.4. Engine efficiency

In order to evaluate LT-SE efficiency taking into account regenerator characteristics, we adopt (7). Pressure drop in other heat exchangers (absorber and cooler) was not considered, as it is relatively low in comparison with the charge loss in the regenerator and does not affect the regenerator evaluation [8][22].

$$\eta = \frac{-(\mathrm{W_{ind}} - \mathrm{W_r})}{-\mathrm{W_c} + \mathrm{Q_{perte}}} \quad (7)$$

W_{ind} is the indicated work, which does not take into account any loss. W_r is the work lost in the regenerator by friction of the fluid against regenerator fibers. W_c is the work during the expansion. And Q_{perte} is the heat that hot source should give additionally, due to the regenerator imperfection.

W_e suppose that during the passage of the fluid from the hot side to the cold side, the same heat amount (Q_{perte}) should be evacuated additionally by the cold source, due to the regenerator imperfection. These energies are calculated as in (8).

$$\mathrm{Q_{perte}} = \mathrm{m C_p}\,(\mathrm{T_c} - \mathrm{T_f})(1-\varepsilon) \quad (8)$$

The indicated work is calculated based on the variation of the total volume and the instantaneous pressure. Two Schmidt hypothesis were adopted; the harmonic variation of engine volumes, and the uniform pressure inside the engine [23]. Our machine can be represented as in Figure 2.

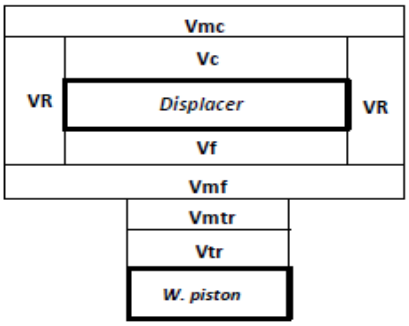


Figure 2: Schematic representation of a beta Stirling engine [23]

V_{mc}, V_{mf} and V_{mtr} are dead volumes in the hot side, cold side and in the working space respectively. V_R is the regenerator volume, V_{tr}, V_C and V_f are the working space, hot and the cold instantaneous volumes respectively.

dV_e and dV_c are respectively the variation of the total volume during the expansion and the compression. V_e, V_c, V_{tr} are defined in (9 - 12).

$$Vc = \frac{V_{sc}}{2}(1 + cos(\theta + \alpha)) \quad (9)$$

$$Vf = \frac{V_{sc}}{2}(1 - cos(\theta + \alpha)) \quad (10)$$

$$V_{tr} = \frac{V_{Str}}{2}\,(1 + cos\theta) \quad (11)$$

Applying the ideal gas law, we can write:

$$p = mr\left(\frac{Vc}{Tc} + \frac{V_f}{T_f} + \frac{V_{Tr}}{T_f} + \frac{Vr}{Tr} + \frac{V_{mc}}{Tc} + \frac{V_{mf}}{T_f} + \frac{V_{mtr}}{T_f}\right)^{-1} \quad (12)$$

Using Schmidt theory for our beta type Stirling engine, we find the following expression of instantaneous pressure, according to engine flywheel rotation angle θ [23]:

$$\mathrm{P} = \frac{\mathrm{mr/k}}{1 + (\frac{\mathrm{A}}{\mathrm{k}})\cos\,(\theta - \beta)} \quad (13)$$

With:

$$k = \frac{V_{SC}}{2Tc} + \frac{V_{SC}}{2Tf} + \frac{V_{Str}}{2Tf} + \frac{2V_r}{Tf + Tc} + \frac{V_{mc}}{Tc} + \frac{V_{mf}}{Tf} + \frac{V_{mtr}}{Tf} \quad (14)$$

$$A = \frac{1}{2}\left[\left(\frac{V_{Str}}{Tf}\right)^2 + \left(\frac{Vsc}{Tc} - \frac{Vsc}{Tf}\right)^2 - 2\frac{V_{Str}}{Tf}\left(\frac{1}{Tf} - \frac{1}{Tc}\right)Vsc \times cos\alpha\right]^{\frac{1}{2}} \quad (15)$$

$$\tan\beta = \frac{\left(\frac{1}{Tf} - \frac{1}{Tc}\right)\sin\alpha}{\frac{V_{Str}}{Tf \times Vsc} - \left(\frac{1}{Tf} - \frac{1}{Tc}\right)\cos\alpha} \quad (16)$$

W_r is calculated based on the pressure drop ΔP [20]:

$$W_r = -\oint \Delta P\, dVc = -\sum \Delta P\, dVc \quad (17)$$

$$W_{ind} = -\int P\, dVc - \int P\, dVf = -\sum P\, dVc - \sum P\, dVf \quad (18)$$

$$W_e = -\int P\, dVc = -\sum P\, dVc \quad (19)$$

3. Validation set up and parametric study

The correlating expressions for pressure drop produced by wood [11] concern two types of regenerator material: woven screens and metal felts. These felts have a random fiber orientation but with the fibers lying predominantly transverse to the flow direction. In SUNWATER 3 prototype, we use random fiber regenerator due to its simplicity and cost. For this reason, an experimental study is carried out to verify the validity of these correlations for our case. Measurement of pressure drop was conducted for SUWATER 3 [24] in order to compare experimental results to theoretical calculations. Pressure sensors in hot and cold sides were installed (Figure 3).

After the experimental validation, effects of many regenerator parameters are studied for finding their optimal values. The parameters are: engine speed, regenerator volume, porosity, wire diameter, working fluid type, and fibers arrangement. The impact of the former parameters on pressure drop in the regenerator, heat transfer coefficient, the regenerator and the engine efficiencies was investigated using equations developed in the theoretical analysis.

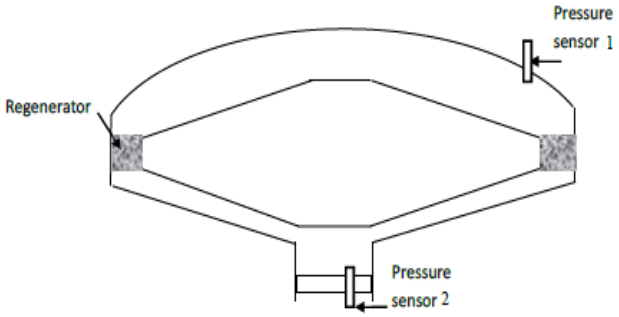


Figure 3: Experimental measurement of pressure drop [24]

Results were found using a calculation program on Microsoft EXCEL, which calculates for a thermodynamic cycle (a complete revolution of the engine flywheel), at a step of 1°, the engine internal chambers instantaneous volumes Vc , Vf and Vtr, the mass flow rate, the instantaneous pressure, the hydraulic diameter, the fluid velocity, Reynolds number, pressure drop, heat transfer factor, etc.

The prototype SUNWATER 3 was taken as a calculation example. Table 2 indicates data for studied cases. Each studied case shows effect of the variation of selected parameter on engine performance.

4. Results and discussions

4.1. Experimental validation of pressure drop correlation

Figure 4 presents experimental results. It shows the variation of the difference between pressures given by sensors 1 and 2 in time. This pressure difference represents the pressure drop in the engine.

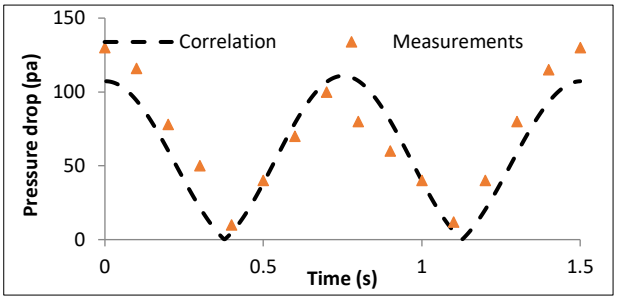


Figure 4: Pressure drop measurement and calculation during one cycle

The comparison between experimental and theoretical curves shows a good agreement, despite some small differences. These differences are due to the structure of our regenerator, and to the fact that sensors give pressure drop caused by regenerator and heat exchangers in hot and cold side, whereas the model gives the pressure drop caused by the regenerator only. Indeed, the pressure drop in the machine is mostly due to the regenerator. This result validates the calculation model adopted, especially for pressure drop correlation. Therefore, we can use the developed model to conduct a parametric study for optimization of LT-SE.

4.2. Effect of engine speed

In order to find the effect of engine speed on regenerator and engine efficiencies, we calculated the former values for many

engine's speeds (data are available in Table 2, case 1). Results are shown in Figure 5.

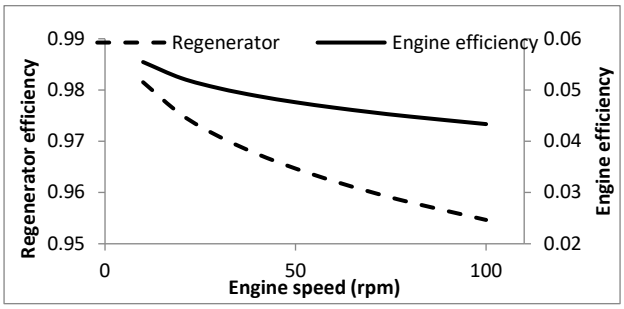


Figure 5: Variation of regenerator and engine efficiencies according to engine speed

Table 2: The parameters range of parametric study

Case	Engine Speed *(rpm)*	Regenerator length *(m)*	Porosity β (-)	Wire diameter *dw* (mm)	Fluid type	Fibers type
1	[10-100]	0.1	0.9	0.2	Air	random
2	40	[0,03-0,5]	0.9	0.2	Air	random
3	40	0.1	[0,6-1]	0.2	Air	random
4	40	0.1	0.75	[0,05-0,5]	Air	random
5	40	0.1	0.8	0.2	variable	random
6	40	0.1	0.8	0.2	Air /Helium	variable

Figure 5 shows that engine performances decrease as engine speed increases. This result is due to the increasing of pressure drop in regenerator with speed. High speed has also bad effect on heat transfer in regenerator, because of the low time contact between the gas and the matrix. Hence, it is advisable to run LT-SE at a low speed in order to improve its performance.

4.3. Effect of regenerator volume

In order to identify the regenerator volume effect on LT-SE for a fixed cross section area, we varied the regenerator length (Table 2, case 2) and calculated pressure drop, convective heat transfer coefficient and regenerator and machine efficiencies. Results are shown in Figures 6 and 7.

Figure 6 indicates that heat transfer coefficient is not affected by the regenerator length, as hydraulic diameter and engine speed have not changed. However, pressure drop decreases with the increase of the regenerator length.

Figure 7 indicates that it exists an optimal regenerator length *(L_{reg}= 0.12m),* for which the engine efficiency is maximum. After this value, the negative effect of pressure drop is more significant than the increasing of the contact surface between the gas and the regenerator. The optimal regenerator length for LT-SE is then about:

$$L_{reg} = 1.41 \times \text{diplacer stroke}$$

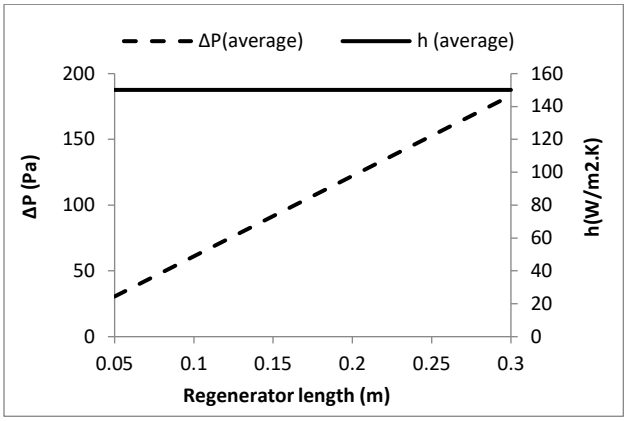


Figure 6: Effect of regenerator length on pressure drop and heat transfer coefficient

4.4. Effect of porosity

Figure 8 presents the effect of porosity on pressure drop and heat transfer coefficient of the regenerator (Table 2, case 3). It indicates that pressure drop decreases with increase of regenerator porosity. Considering heat transfer coefficient, the value of porosity which maximizes it is β=0.8.

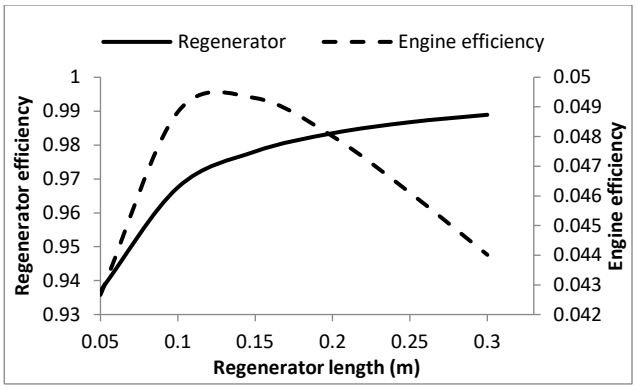


Figure 7: Effect of regenerator length on engine and regenerator efficiencies

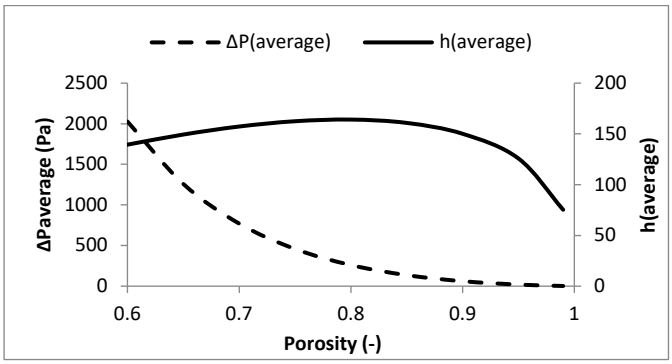


Figure 8: Variation of heat transfer coefficient and charge loss according to porosity

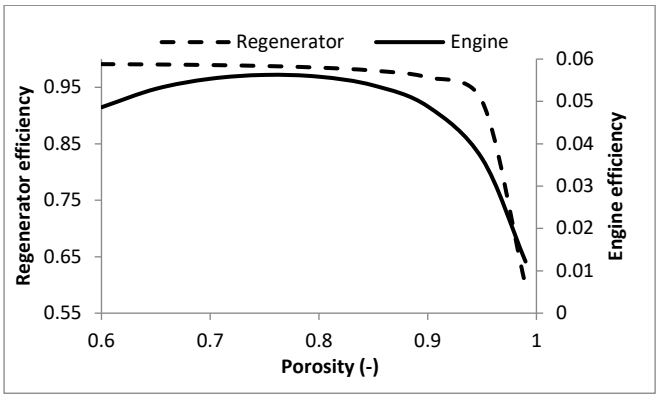


Figure 9: Variation of regenerator and engine efficiencies according to porosity

Figure 9 shows the effect of porosity on the regenerator and engine efficiencies. It indicates that these efficiencies follow the same trend as the heat transfer coefficient. Thus, it's recommended to adopt the optimal value $\beta=0.8$, giving the best performances.

4.5. Effect of regenerator wire diameter

The effect of regenerator wire diameter (Table 2, case 4) is illustrated on Figure 10. An optimal value for wire diameter is found to be d_w=0.*15mm*, corresponding to hydraulic diameter d_h=0.*45mm.*

Work of Tanaka et al. [8] based on experimental tests indicates also an optimal wire diameter. They found, for regenerators longer than 70mm (which is the case for our adopted values), that there is an optimal value for wire diameter. But it was not defined, due to smaller variation range of regenerator length.

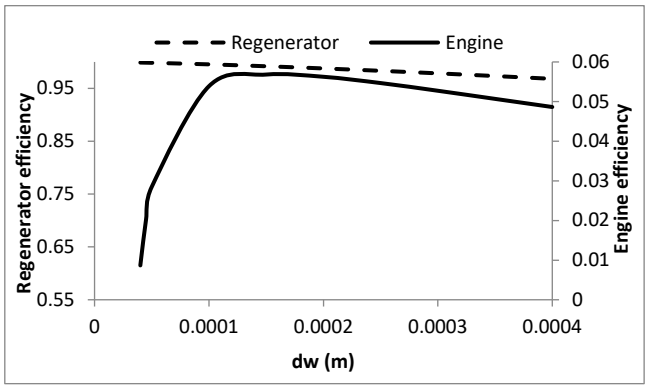


Figure 10: Variation of engine and regenerator efficiencies according to wire diameter dw

4.6. Effect of working fluid

Effects of three working fluids: hydrogen, helium and air, on regenerator and engine efficiencies were investigated.

Figures 11 and 12 present the effect of working fluid type on the variation of heat transfer coefficient and pressure drop during a complete cycle. Calculation data are available in Table 2, case 5. These figures indicate that hydrogen is the most performant fluid among the studied fluids, giving the best heat transfer coefficient and causing the less pressure losses. These results are due to the fact that hydrogen is the lightest and less viscous fluid among the studied fluids (see Table 3).

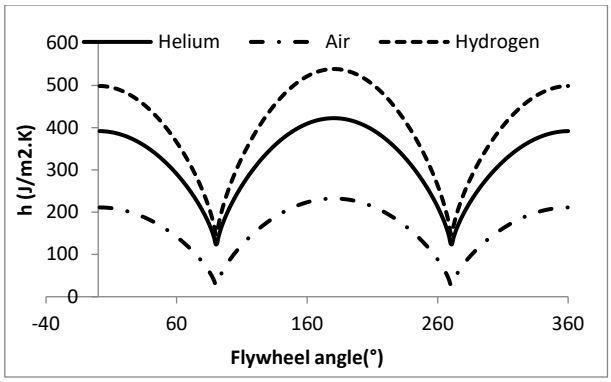


Figure 11: Cyclic variation of heat transfer coefficient for several working fluids

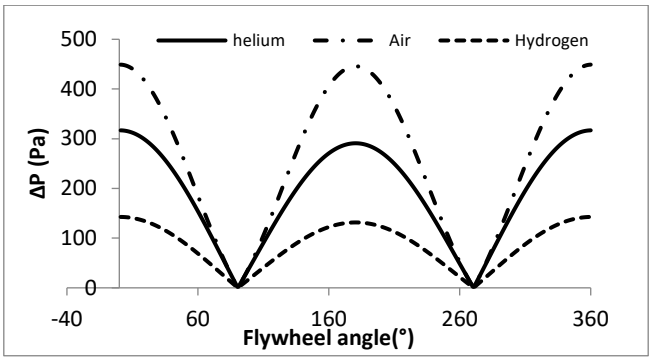


Figure 12: Cyclic pressure drop variation for several working fluids

Figure 13 indicates the regenerator and engine efficiencies for the studied working fluids. Helium and hydrogen give the best performance compared to air. We can also note that although hydrogen has the highest heat transfer coefficient and causes less pressure drop, helium gives the best engine performances.
This is due to the fact that helium density is greater than that of hydrogen (see Table 3), which has the effect of increasing the mass of gas inside the machine, and causes consequently more energy absorption.

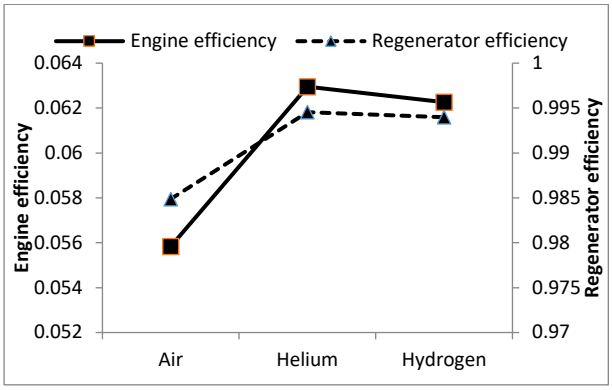


Figure 13: Engine and regenerator efficiency for many working fluids

Table 3: Fluids properties [25]

Fluid	Density at *(20°C)*	Viscosity at *20°C* (Pa.s)	Specific Heat *Cp* at *20°C (J/Kg K)*
H_2	0.0827	8.7454E-06	14307
Helium	0.16422	0.000019586	5195
Air	1.204	0.00001825	1007

4.7. Effect of regenerator type

Figure 14 shows the effect of the regenerator fibers arrangement on engine efficiency (Table 2, case 6). Studies were made for air and helium due to availability of air and high performance of helium.

This figure indicates a higher performance for woven screens regenerators in comparison with random fiber ones, especially for speeds greater than 50 rpm. However, the speed range of LT-SE is generally low, as confirmed by Kolin [26]. Then, while designing LT-SE, the use of random fibers is a justified choice due to the acceptable performance, availability and low price.

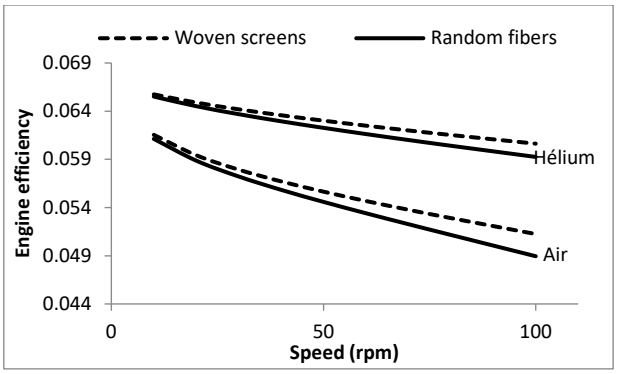


Figure 14: Engine efficiency for two regenerator types

5. Conclusion

In order to contribute to the development of LT-SE regenerators' design, a calculation model was developed, validated and used to analyze the effect of six regenerator parameters in order to find out their optimized values. Studied parameters are: engine speed, regenerator volume, regenerator porosity, regenerator wire diameter, working fluid type and regenerator fibers arrangement. Results indicated a number of optimization recommendations of LT-SE:

- Low speed is preferred in order to reduce the pressure drop and to enhance the heat exchange in regenerator;
- A regenerator length around L_{reg} should be adopted, where :

$$L_{reg} = 1.41 \times diplacer\ stroke$$

- The optimal value for wire diameter is: $d_w = 0.15mm$, corresponding to a hydraulic diameter $d_h = 0.45mm$;
- Porosity of the regenerator matrix should have a moderate value $(\beta = 0.75)$;
- The most performant working fluid is helium. However, because of the recurring sealing problems on Stirling engine performance, the use of air is most adequate (available and for free);
- The use of woven screens regenerators is not necessary for engines running at low speed *(<50rpm),* random fibers give acceptable performance. At higher speed, woven screens regenerators are more suitable, giving better engine efficiency.

By following the guidelines presented in this paper, an appropriate configuration for LT-SE can be selected.

Nomenclature

Parameter	Signification- Unit
A	Section area of the regenerator housing (m^2)
A_{wg}	Wetted area (m^2)
Cp	Specific heat of fluid at constant pressure $(J/kg\ K)$
d_h	Hydraulic diameter (m)
d_w	Wire diameter (m)
f	Darcy Friction factor (-)
h	Convective heat transfer coefficient $(W/m^2\ k)$
k	Thermal conductivity (W/mK)

L_{reg}	Regenerator length (m)
Nu	Nusselt number (-)
NTU	Number of transfer units (-)
Pe	Peclet number (-)
Pr	Prandtl number (-)
Q_m	Mass flow rate (kg/s)
Q_{perte}	Heat lost due to regenerator imperfection (J)
Re	Reynolds number (-)
r	Specific gas constant (*J/kg K*)
T_c, T_f	Absorber a cooler temperatures respectively
Tr	Regenerator temperature
u	Fluid velocity (*m/s)*
V_c	Expansion volume
Vf	Compression volume
V_{mc}	Dead volumes in the absorber side
V_{mf}	Dead volume in the cooler side
V_{mtr}	Dead volume in the working space
V_{sc}	Maximal volume of hot space
V_{str}	Maximal volume of working space
V_R	Regenerator volume
V_{tr}	Working space instantenous volume
Greek symbols	
α	Phase angle (°)
β	Porosity (-)
θ	Flywheel angle (°)
ρ	Density (*kg/m3*)
ε	Regenerator efficiency (-)
μ	Dynamic viscosity (*kg/ms*)
η	Engine efficiency (-)

Conflict of interest

The authors declare no conflict of interest.

References

[1] R.Stirling, Stirling air engine and the heat regenerator, Patent no, 4081, 1816.

[2] K.Wang,S.R. Sanders, S. Dubey, F.H. Choo, F. Duan, " Stirling cycle engines for recovering low and moderate temperature heat: A review" Renew. Sustain. Energy Rev., 62, 89–108, 2016. https://doi.org/10.1016/j.rser.2016.04.031

[3] C. Toro, N. Lior, "Analysis and comparison of solar-heat driven Stirling, Brayton and Rankine cycles for space power generation" Energy, 120, 549–564, 2017. https://doi.org/10.1016/j.energy.2016.11.104

[4] S K. Andersen, H. Carlsen, P.G. Thomsen, "Numerical study on optimal Stirling engine regenerator matrix designs taking into account the effects of matrix temperature oscillations" Energy Convers. Manag. , 47, 894–908, 2006. https://doi.org/10.1016/j.enconman.2005.06.006

[5] L. Sun, T.w. Simon, S. Mantell, M. Ibrahim, D. Gedeon, R. Tew, "Thermo-fluid experiments supporting microfabricated regenerator development for a Stirling spacepower engine", in 2009, 7th International energy conversion engineering conference, Denver, Colorado 2009. 10.2514/6.2009-4579

[6] W. Kays, A. London, Compact Heat Exchangers.,New York: McGraw-Hill, 1984.

[7] J.R. Sodré, J.A. R. Parise, "Friction Factor Determination for Flow Through Finite Wire-Mesh Woven-Screen Matrices" J. Fluids Eng., 119, 847–851, 1997. https://doi.org/10.1115/1.2819507

[8] M. Tanaka, I. Yamashita, F. Chisaka,"Flow and heat transfer characteristics of the stirling engine regenerator in an oscillating Flow" JSME Int. J., 33, 283–9, 1990. https://doi.org/10.1299/jsmeb1988.33.2_283

[9] S. Alfarawi, R. Al Dadah, S. Mahmoud, "Potentiality of new miniature-channels Stirling regenerator", Energy convers. Manag., 133-264, 2017. https://doi.org/10.1016/j.enconman.2016.12.017

[10] I. Rühlich, H. Quack , New regenerator design for cryocoolers, Technische Universiat, Dresden,1999.

[11] D. Gedeon, J. Wood, Oscillating-flow regenerator test rig: hardware and theory with derived correlations for screens and felts, NASA CR (N°198442), 1996.

[12] MB. Ibrahim, T. Simon, D. Gedeon, R. Tew , "Improving performance of the Stirling converter: redesign of the regenerator with experiments", Computation and modern fabrication techniques, Cleveland State University, 2001.

[13] MB. Ibrahim., Danila, D., Simon, T.W., Gedeon, D., Tew, R. , "Computational Modeling of a Segmented-Involute-Foil Regenerator for Stirling Engines" J. Thermophys. Heat Transf., 23, 786–800, 2009. https://doi.org/10.2514/1.40330

[14] R. Tew , R. Dyson, S.D. Wilson, R. Demko. , Stirling convertor CFD model development and regenerator R&D efforts, in December 2004, In 2004, AIP Conference Proceedings 746,NASA, 2004. DOI: 10.1063/1.1867183

[15] Anders S. Nielsen, Brayden T. York, Brendan D. MacDonald, " Stirling engine regenerators: How to attain over 95% regenerator effectiveness with sub-regenerator and thermal mass ratios", 253, Applied Energy, 2019. https://doi.org/10.1016/j.apenergy.2019.113557

[16] K.V. Srinivasan, A. Manimaran, M. Arulprakasajothi, M.Revanth, Vijay A. Arolkar, "Design and development of porous regenerator for Stirling cryocooler using additive manufacturing" Therm. Sci. Eng. Prog. , 11, 195-203, 2019. https://doi.org/10.1016/j.tsep.2019.03.013

[17] N. Boutammachte, J. Knorr, "Field-test of a solar low delta-T Stirling engine" Sol. Energy, 86, 1849–1856, 2012. https://doi.org/10.1016/j.solener.2012.03.001

[18] H. El Hassani, N. Boutammachte, M. Hannaoui, "Study of some power influencing parameters of a solar low temperature Stirling engine", European Journal of Sustainable Development, 3(2), 109-118. http://dx.doi.org/10.14207/ejsd.2014.v3n2p109, 2014.

[19] R. Gheith, F. Aloui , S. Ben Nasrallah, "Study of temperature distribution in a Stirling engine regenerator", Energy Convers. Manag., 88, 962-972, 2014. https://doi.org/10.1016/j.enconman.2014.09.043

[20] K. Hirata, I. Iwamoto, K. Hamaguchi, "Performance evaluation for a 100 W Stirling engine" in 1997 8th International Stirling engine conference, University of Ancona, Italy, 19–28, 1997.

[21] S. Wilson, R. Dyson, R. Tew, R. Demko, Experimental and computational analysis of unidirectional flow through Stirling engine heater head". Nasa center for aerospace information, 2006.

[22] D. Berchowitz, I. Urieli, Stirling Cycle Machine Analysis. Adam Hilger, 1984.

[23] H. El Hassani, N. Boutammachte, J.Knorr, M. Hannaoui, "Study of a low-temperature Stirling engine drivenby a rhombic drive mechanism", Int. J. Energy Environ. Eng., 4, 40-50, 2013. https://doi.org/10.1186/2251-6832-4-40

[24] D. Chen, "Untersuchung zur Optimierung eines solaren Niedertemperatur-Stirlingmotors", Ph.D Thesis, Technischen Universität Dresden, 2004.

[25] H. El Hassani, " Etude théorique et expérimentale des paramètres régissant le fonctionnement des moteurs Stirling à basse temperature", Ph. D Thesis, Moulay Ismail University, 2014.

[26] I. Kolin, Stirling Motor: History- theory- practice, Dubrovnik- Zagreb University Publications, 1991.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 275-283 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Angular Orientation of Steering Wheel for Differential Drive

Rajesh Kannan Megalingam*, Deepak Nagalla, Ravi Kiran Pasumarthi, Vamsi Gontu, Phanindra Kumar Allada

Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, 690525, India

ARTICLE INFO

Article history:
Received: 15 January, 2020
Accepted: 29 April, 2020
Online: 29 May, 2020

Keywords:
Differential drive
Robot Operating System (ROS)
Wheeled Robot
Independent Control
Angular Motion
Gazebo
Universal Robot Description Format (URDF)
ROS_Node
Trajectory

ABSTRACT

Several drive mechanisms for different robots are at hand in current days. Bicycle steering, Ackerman steering, differential drive are some principal drive mechanisms that are being deployed in robots these days. The differential drive needs the wheel rotations to be updated very frequently. But it is most commonly deployed on the robots with two wheels and casters, as discussed in this work. It also can be used to have an independent control for each of the wheels with independent control signals. This work deals with the modeling of the differential drive mechanism for a robot with two main drive wheels and two casters, which takes the angular orientation of the steering wheel as input. For simplicity, this work considers that casters have no influence on any aspect of the differential drive. An adaptive model, whose output depends on the real-time input from the gamers steering wheel and produces required output has been formulated. This work differs from the other differential drives in the context that the steering wheel and the robot wheels have no physical connection. The proposed model has been implemented in python and integrated with the Robot Operating System (ROS). The steering wheel, which is used to generate commands using, is attached to the controller at the control station and the ROS_Node thus created is used to read the values from the steering and generate commands for each of the left and right wheels. These commands are transferred to the controller on the mobile platform, which in turn generates control signals for actuators. This work also deals with the deployment of the proposed model using the Universal Robot Description Format (URDF) of the robot in the Gazebo simulation and evaluating it using the Nitho Drive Pro steering wheel. To prove that the differential drive mechanism can be used to control the robot efficiently in any type of terrain, a ROS python node is used to control and maneuver the robot.

1. Introduction

Mobile robots are nowadays being deployed in varied scenarios. Different designs are needed in dissimilar scenarios for effective deployment and efficient use. Each design will be based on a unique drive mechanism [1], which best suits the arena for the implementation of the robot. Some of the most commonly used steering mechanisms are Mecanum drive, Synchronous drive, XR4000, Differential drive and Ackerman steering. Each mechanism has its own not only pros but also cons, which can sometimes be advantageous during installation in the area of interest. Differential drive, which is of interest in this work, works based on the difference in angular velocities of the two wheels of the robot as in work [2]. It has its own disadvantage in requiring the velocities to be updated very frequently. But this mechanism best suits the requirement when the wheels should be operated independently as stated in paper [3]. Especially when the wheels or their alignment is not identical, they should be controlled independently for stability and better control. The robots, though designed based on a drive mechanism, need the control signals to maneuver. These signals can be either autonomously generated or from an HMI device as discussed in paper [4]. Gamer's steering wheel is one such device which consumes low power and gives an interactive control experience to the user. This can be used to generate the required commands for the differential drive. Unlike implementing this mechanism in cars, most of the robots are wireless and the user cannot be on the robot in all instances. The

*Rajesh Kannan Megalingam, Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India, 09496120900 & rakeshkannan@ieee.org

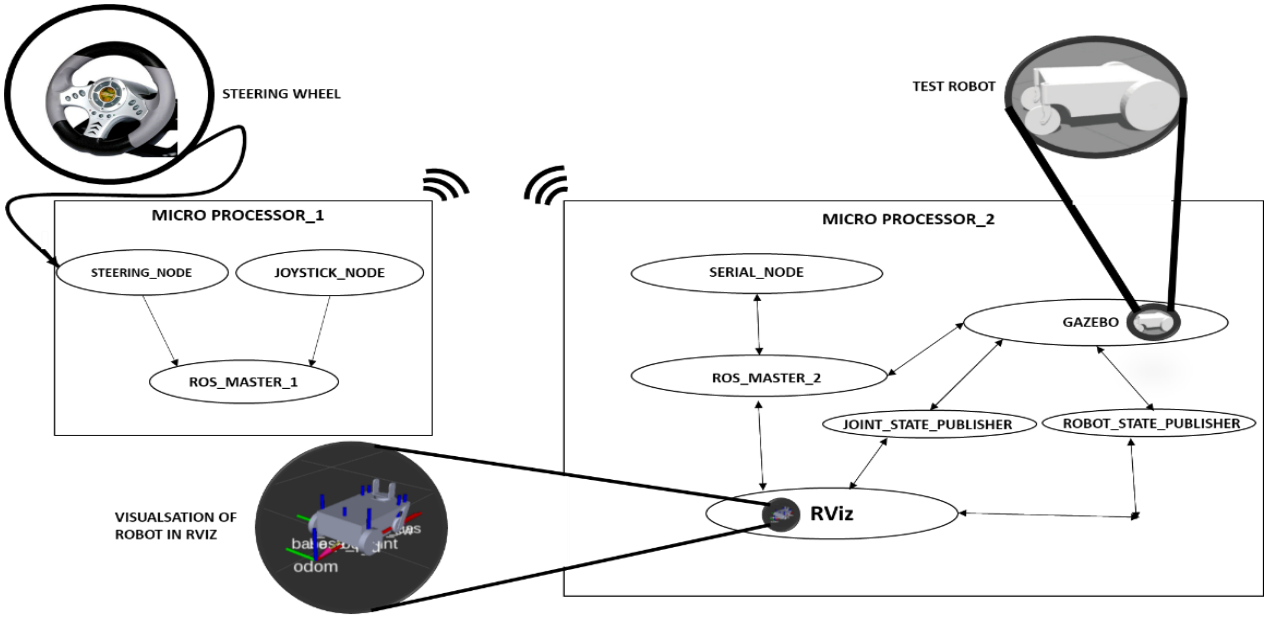


Figure 1(a): Internal ROS communication

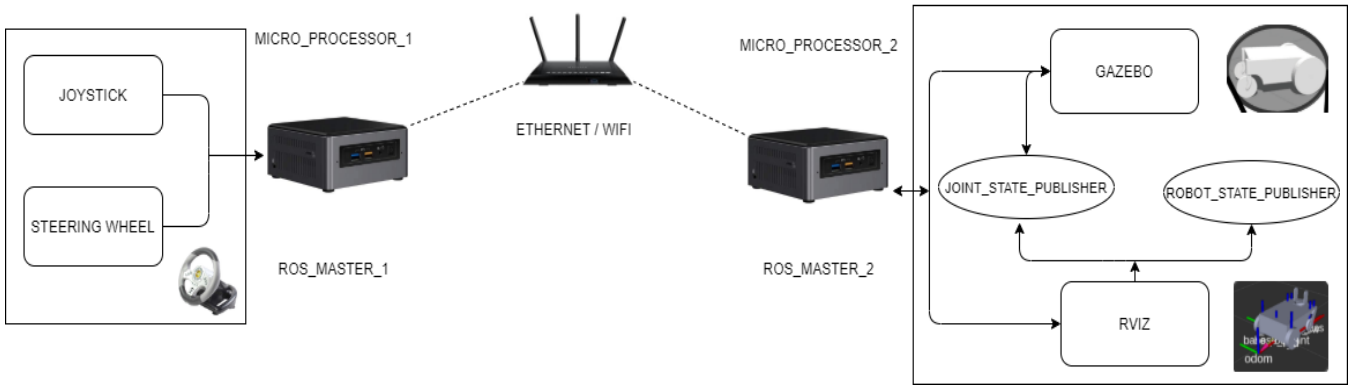


Figure 1(b): Architectural diagram

robot might be at one place and the HMI can be at another. ROS comes handy in such scenarios, especially when data should be exchanged between different nodes, which might be on different systems. ROS acts as a strong platform with a variety of tools and tremendous support for any robotics application. Gazebo, a simulation tool, which is interfaced with ROS as described in the work [5] and can be used to simulate mostly all kinds of robots with simulated arenas, which resemble the real-time environments. And the simulation is plain sailing, a robot or an environment can be built directly in the simulation tool or a URDF file generated from solid works can be used for simulation.

2. Motivation and Problem Statement

Considering the best example for this case is cars with power steering gives an appropriate turn without using much power on the steering. This theory of power steering helps to avoid accidents in vehicles by reducing the effort on the driver and with exact turnings. Inaccurate performance of the vehicle highlights neither the model of the robot nor the working. Many gamers might have used this steering wheel in different PlayStation games in order to get the exact sensation of the game. Comparing this robot with the cars in the games like F1 race, blur and many. Considering the same case in the wheeled robots, implementing this idea may give an appropriate solution. To overcome this problem the following concepts have been proposed. The concept and the modeling [6] in this paper mainly concerned the differential drive with two castor wheels (freewheels) [7], and two main actuators [8]. This model helps the user to control from a remote area where there is no point in direct contact with the robot. The ROS-Gazebo [9] interface with this robot gives the simulation measures that help in achieving accurate results.

3. Related Works

Paper [10] mainly deals with the Varying speeds of differential drive motors and for no of stages. This approach of considering the varying speeds, which is a basic case in any differential drive approach has been adapted. This work deals with a better way of handling highly varying speeds in different scenarios.Kinematic algorithms and mobility control algorithms are explained in [11]. High reliability and precise working mechanisms with powerful dynamic equations can be taken from [12].The system modeling of a four-wheeled differential drive

robot which used to achieve accuracy in Control by using sliding mode control is proposed in [13].A hierarchical coordination control approach by defining various layers has been taken from [14].Designing and simulating the robot, using the solid works, nonlinear dynamics to control the simulated robot is presented in [15].The algorithm of the simulation control of robots and the behavior of the robot during simulation is presented in [16]. Working principle of controllers and the new form of feedback to controllers, an easy understanding of the control system is presented in [17]. Designing our own form of the controller and building the interface between the controller and the robot platform is presented in [18]. Wireless communication between the master controller and the slave robots, the algorithm for a wheeled robot is presented in [19]. The differential drive motor control technique has been taken from learning robotics using python as presented in [20]. Paper [21] explains the motion planning of the robot using the differential drive and trajectory smoothing using optimization techniques.Mathematical modeling and behavior of the robot are explained using the skid steering model, the tracking of the robot is also mentioned in paper [22].

4. Architectural Diagram

Figure 1(a) shows the ROS connections and communications of the entire system and how it controls the robot from a control station to the remote position of the robot. The left block in the diagram i.e. with the steering wheel explains the internal process happening at the control station. Similarly, the right block explains the internal connections and the processes happening within the robot. Figure 1(b) shows the architectural block diagram of this robot.

4.1. ROS_Master_1

The master controller of the MICRO_PROCESSOR_1 is entirely dependent on the ROS_MASTER_1 node which is independent of the entire package once the roscore is launched it automatically starts. Once the ROS_MASTER_1 is activated it acts as a host and starts connecting to different nodes (or vice versa) with different topics as a bridge and communicating among themselves. This node runs in the control station and communicates within the nodes i.e.; STEERING_NODE and JOYSTICK_NODE.

4.2. Steering_Node

The input to the ROS_MASTER_1 is given from either keyboard or steering wheel, which is connected to the processor through I/O peripherals. Keys like ‘i’, ‘,’, ‘j’, ‘l’ in the keyboard are used for the movement of the robot in forward, backward, right, left respectively. Speed control can also be done using the keyboard. The steering wheel can be used for the same and the control resembles driving a car. These devices' steering wheel, joystick should be integrated and accessed with the ROS_MASTER_1 through STEERING_NODE and JOYSTICK_NODE.

4.3. Joystick_Node:

These nodes form a bridge between I/O devices and ROS. Each key of these devices has a unique ID in these nodes. So, when a key is pressed, the respective ID gets active and this, in turn, can be assigned with a function to be performed. The values or commands generated in this function are published over a topic that can be accessed by the Gazebo node, where the robot is being simulated in different environments.

4.4. ROS_Master_2

Microcontroller_2 has another ROSmaster named ROS_MASTER_2 which is in the mobile robot that communicates with the nodes at the remote area i.e. within the robot. Since the entire communication of the system is done through wireless connectivity, both the ROS_MASTER’s are controlled using the master-slave protocol.

4.5. Gazebo

The gazebo is open-source for the simulation of different environments. Gazebo interfaces with ROS and communication are done using the GAZEBO node. It can act as a publisher as well as a subscriber. GAZEBO node subscribes to the commands published by the STEERING_NODE and JOYSTICK_NODE to move the joints of the robot accordingly. It also publishes the transforms of the URDF node which in turn helps to move the robot. The Joint_states of the robot’s movement is published (JOINT_STATE_PUBLISHER)over the Odom topic which is used by the differential drive plugin to correct the robot’s path from the intended path. This data is also used by RViz to mimic the robot in Gazebo.

4.6. Joint_State_Publisher

This is used to read data from various origins. This has special features such as inputting the values through the graphical user interface, sending those values to the other joint and can set the default values even. It includes various parameters such as robot description, velocities, position, dependent joints...etc.

4.7. Robot_State_Publisher

ROBOT_STATE_PUBLISHER is used to publish the state of the robot to a tool where it can keep track of multiple coordinate frames per overtime. For a package, the position of the joint is taken as the initial values. This package subscribes to the joint states of the robot and publishes a three-dimensional representation of each and every link using the kinematic involved in the URDF model. Internally it has a kinematic model of the robot. General parameters used by this node are robot description, tf_prefix, publish frequency and it subscribes to the node joint states.

4.8. RViz

RViz is a platform for visualization and displaying different sensor values and information using ROS. It helps us to analyze different aspects of the robot-like transforms.

Transforms play a dominant role in avoiding collisions and in establishing coordination amongst the different frames and regulating the movement of all the joints. Not just the transforms of physical systems but also with the odometry frame can be calculated. This is not a simulation tool but a visualization interface where we can visualize both simulated robots and a real-time robot as shown in Figure 2.

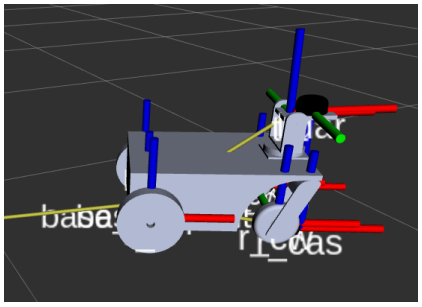


Figure 2: Robot model in RViz with TF's

4.9. Serial_Node

SERIAL_NODE introduces serial communication which is a bit to bit transmission of data. Rosserial helps this type of communication in different hardware parts of microcontrollers such as Arduino, teensy and also acts as an interface between the hardware devices and software tools.

5. Design and implementation

5.1. Implementation in a software platform

Nodes are a part of ROS in which they communicate with the other nodes in order to transfer the data or information. These nodes are generally connected by a bridge called a topic. The topic is a unidirectional data channel. The communication between these nodes and topics can be viewed through graphs(Rqt_graph) or frames.

The simulator consists of launch and configuration files. A virtual robot can be built in Gazebo which produces a xacro file that has all the links connected to the parent link. Similar to the Gazebo tool we have many other platforms (v-rep, webots, solid works) in which Solid works are most prominent for designing.

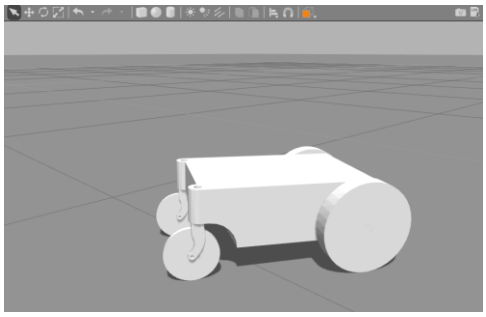


Figure 3: Gazebo Bot

Solid works to URDF exporter has been used to generate a URDF file from the solid works model, which includes inertia on the body, torque on the wheel links, parent links, velocity as shown in Figure 3. For the obtained URDF file adding the differential-drive [23] motor controls the velocity of the wheels independently. Forward kinematics equations are used in differential-drive to solve the problems obtained during the slippery surfaces controlling velocity.

$$V = r\omega \quad (1)$$

where V is the velocity of the wheel, r is the radius of the wheel, $\omega$is the angular velocity of the wheel.

This above Equation (1) is the basic fundamental for this differential-drive. This URDF file is added into a ROS package and connected to the master and we can navigate the robot in any environment. Once the body is computationally ready with all the connections within the code, the controller launches all the required files using ROS commands. These launch files communicate through the nodes to publish or access the required data. The topic of Joint_states published by the Gazebo node and subscribed to the keyboard node using the command velocity(/cmd_vel) topic.

5.2. Steering wheel control

Nitho Drive Pro is a Steering wheel compatible with PC2, PC3 and PC. It helps us to control the Wheeled robot using teleoperation. In the steering wheel, the Joint_state_values are published in the terminal, which is in turn subscribed by the robot, which helps in teleoperation. This robot simulation helps us to get the virtual feedback in Gazebo, so the operator knows what is happening just by looking in the simulation. By using the steering wheel, we can even adjust the speed and also the direction we want to turn using axis values. We can assign different values for buttons ranging from 0 to 1. We can change the values of the axes to turn based on our needs. The Steering Wheel system consists of two different controls for the forward and backward motion. One is pedal motion and the other is by using the buttons on the steering wheel.

This Steering wheel system can be implemented in an actual car driving for handicapped people who cannot apply much force to turn the wheel. Nitho Drive Pro does not require much force to rotate its wheel and it can be programmed in such a way that we can increase or decrease the speed by pressing the designated buttons. We can also find out which buttons are assigned for what number and its functions. By giving the command we can access the control system of the steering wheel and find out the mapping of the steering wheel and be able to change the calibration of the steering wheel. In the mapping option of the steering wheel, we can find out which button is assigned to which number and helps to change the controls accordingly. In calibration default values are assigned to the buttons and axes for movement, we can change the calibration values manually or by altering the values. We can even change the mode of controlling the axes. If the PS/MODE button is on it means that the steering wheel gives the continuous values of the axes, else if it is in the steady mode it means that the axes are in button state, it behaves just like a button and gives a single value. If it is off that means that the steering wheel is off.

5.3. Four-wheeled Differential drive

For the design of a Two-wheeled differential drive model initially, a Four-Wheeled drive model is discussed. This drive system is mainly composed of control and drive systems. The mechanical structure is the basis for the entire robot. It turns using all the four wheels, unlike normal vehicles. Each wheel has a separate motor mechanism to drive the steering of the wheel. Therefore, the motion of all the wheels of the robot are independent of each other.

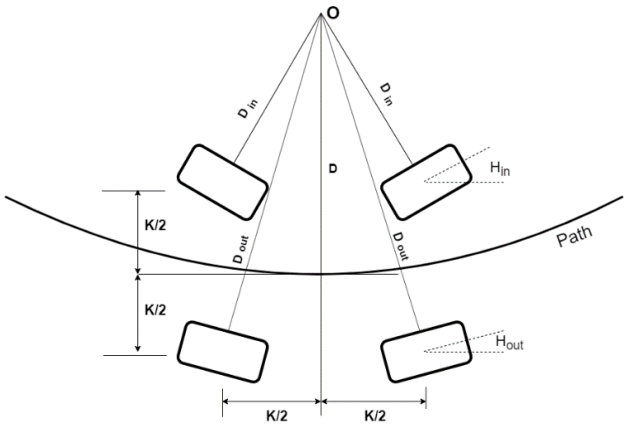


Figure 4: four-wheeled differential drive robot.

From geometric relation, from (Figure 4) it can be considered as follows

$$D = \frac{K(1+\tan(Hin))}{2} \quad (2)$$

$$D_{in} = ((D - \frac{K}{2})^2 + \left(\frac{K}{2}\right)^2)^{\frac{1}{2}} \quad (3)$$

$$D_{out} = ((D + \frac{K}{2})^2 + \left(\frac{K}{2}\right)^2)^{\frac{1}{2}} \quad (4)$$

$$H_{in} = arctan\left(\frac{K}{2D+K}\right) \quad (5)$$

$$H_{out} = arctan\left(\frac{K}{2D-K}\right) \quad (6)$$

Where *D* is the radius of the body in which it rotates, *Din* is the radius of inner wheels in which it rotates, *Dout* is the radius of outer wheels in which it rotates, *K* is the separation between left and right wheels/separation between front and back wheel, H_{out} is the slip angle of outer wheels, Hin is the slip angle of inner wheels. The speed relationship is as follows:

$$s_{in} = \frac{D_{in}S_n}{D} \quad (7)$$

$$s_{out} = \frac{D_{out}S_n}{D} \quad (8)$$

Where s_{in} is the speed of the inner wheels, s_{out} is the speed of the outer wheels, sin is the speed of the robot

$$\psi_{if} = \kappa_{if}\beta_{if} \quad (9)$$

$$\psi_{ir} = \kappa_{ir}\beta_{ir} \quad (10)$$

where i=L refers left, i=R refers right. ψ_{if}andψ_{ir} denote the lateral force of front wheels and rear wheels respectively. κ_{if} and κ_{ir} are the lateral stiffness of the front wheels and rear wheels respectively.$\beta_{if}$is the slip angle for front wheels and $\beta_{ir}$is the slip angle for rear wheels.

5.4. Implementation of a two-wheeled robot

Implementing a robot with a simple 2 wheeled mechanism which has a common axis of rotation. And two caster wheels in the front so there will be no slip angle for the wheels. There will be only one inner wheel and one outer wheel. The moment of the body in circular turns are modeled by the varying speeds of the two wheels. from (Figure 6) the geometric relations of the two-wheeled robots are as follows.

$$D_{in} = D - \frac{K}{2} \quad (11)$$

$$D_{out} = D + \frac{K}{2} \quad (12)$$

$$s_{in} = \frac{D_{in}S_n}{D} \quad (13)$$

$$s_{out} = \frac{D_{out}S_n}{D} \quad (14)$$

$$S_n = \frac{S_{in}+S_{out}}{2} \quad (15)$$

where D is the radius of the body in which it rotates, D_{in} is the radius of inner wheels in which it rotates, $D_{out}$is the radius of outer wheels in which it rotates, K is the separation between left and right wheels/separation between front and back wheels, $s_{in}$is the speed of the inner wheels, $s_{out}$is the speed of the outer wheels, $s_n$is the speed of the robot.

Varying Speeds of front and back motors in terms of rotations per minute.

$$R_{in} = \frac{D_{in}\rho}{D} \quad (16)$$

$$R_{out} = \frac{D_{out}\rho}{D} \quad (17)$$

$$\rho = \frac{[s_{in}-s_{out}]}{k} \quad (18)$$

where R_{in} is the speed of the inner wheel (RpM), ρ is the angular velocity of the robot.

The D (Radius of robot rotation) value will be defined by the steering control as shown in the following graph.

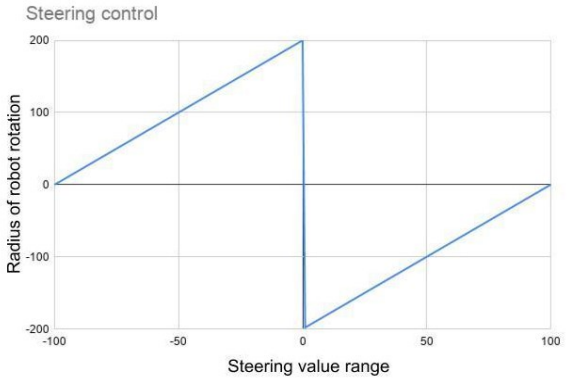


Figure 5: The varying values of D (Radius of robot rotation) according to the steering control (steering readings).

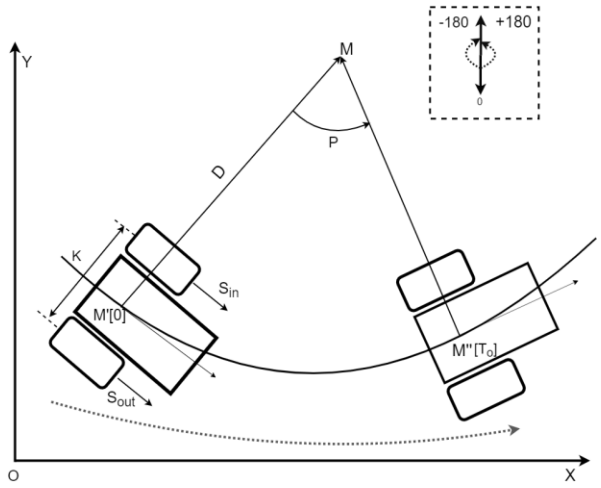


Figure 6: The transformation of the body in a 2D frame.

The robot started at M' at time t=0, the coordinates of the robot at time t=T_o. is M'' the transformation equations are:

$$M = [\, x - Dsin(M'_\theta), y + Dcos(M'_\theta)] \quad (19)$$

$$M''_x = (\cos(\rho T_o))(M'_x - M_x) - (\sin(\rho T_o))(M'_y - M_y) + M_x \quad (20)$$

$$M''_y = (\sin(\rho T_o))(M'_x - M_x) + (\cos(\rho T_o))(M'_y - M_y) + M_y \quad (21)$$

$$M''_\theta = M'_{\theta + \rho T_o} \quad (22)$$

where $[M'_x, M'_y, M'_\theta]$are the coordinates of robot and angle of orientation at M', $[M''_x, M''_y, M''_\theta]$are the coordinates of robot and angle of orientation at M', $[M'_x, M'_y, 0]$are the coordinates of the virtual center of rotation M

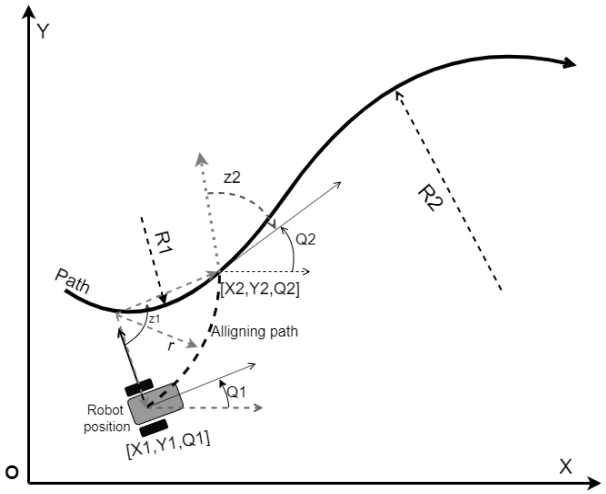


Figure 7: Tracking of the robot according to the steering control values.

The robot is at a random position [X1, Y1, Q1], in order to follow the path on which it should follow it should first align with the path along the aligning path which is mentioned in Figure 7. The aligning path of radius 'r' which meets at a point on the path to be followed after z1 degrees of rotation. The robot now should rotate the z2 angle for making it tangential to the path and follow the path with the radius R2. This path can be tracked in the simulation when controlled using the steering wheel. The following equations give the position of the robot at a time T_o. obtained from the transformation equations.

$$M''_x = \left(\frac{1}{2}\right) \int_0^{T_o} [s_{out}(T_o) + s_{in}(T_o)] \cos(M'_\theta T_o)) \, dt \quad (23)$$

$$M''_y = \left(\frac{1}{2}\right) \int_0^{T_o} [s_{out}(T_o) + s_{in}(T_o)] \sin(M'_\theta T_o) \, dt \quad (24)$$

$$\int M''_\theta = \left(\frac{1}{k}\right) \int_0^{T_o} [s_{out}(T_o) - s_{in}(T_o)] \, dt \quad (25)$$

6. Experiments

To get the real experience of controlling a robot, the Nitho Pro drive steering wheel has been used as shown in Figure 8, to control the robot which is compatible with any type of system. The robot is controlled in Gazebo to get a clear idea of real-life scenarios. Using the differential drive plugin, the angle of rotation of the robot is adjusted to the turning ratio of the steering wheel to get a better idea of the testing scenario and for easy control of the robot as mentioned in [6]. For the controlling of the robot in multiple scenarios using Gazebo, a directory was created. This directory was used to control the robot using a keyboard. After successfully controlling using the keyboard a directory to control using the Steering wheel has been created.



Figure 8: Steering Wheel Contro

Control of the robot using the Steering wheel in gazebo has been successfully implemented using ROS. Differential drive plugin has been added to the URDF file for the control of the robot. The robot has been tested on multiple terrains to test the control and manoeuvrability in different conditions. It was designed in such a way that the speed and angle of rotation of the robot can be monitored from (1), (2) and (3). If the power is kept constant the forces acting on the robot will differ due to the terrain conditions, stiffness, slipping and other forces acting in the body as shown in (9) and (10).

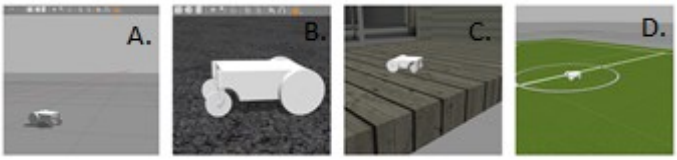


Figure 9: A. Plane surface, B. Asphalt Plane, C. Rough surface, D. Grass field

Multiple testing scenarios were considered for the implementation of the robot model. For the case of flat surfaces in Figure 9.A-D, is considered in many daily life scenarios. Figure A. is a Plane smooth surface where frictional value is zero which cannot be found in our daily life. When the robot is given even a small amount of speed the body moves freely as there is no opposing force but for cases such as Figure 9.B-D, which can be found in our daily a certain amount of force needs to be applied to overcome the frictional force and start moving. In the case of grass fields, the body moves freely due to less frictional force and covers a certain distance in less amount of time as shown in Table.

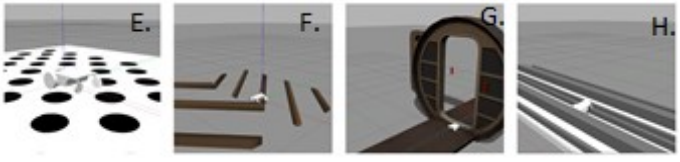


Figure 10: E. Crater surface, F. Uneven path, G. Doorway, H. Moat plane

In the case of scenarios such as rough surfaces in Figure 10..E-H., more speed is required to move over the obstacles but as the power is kept constant the robot moves at a slower speed compared to the plane surface and as a result, it takes more time. For the crater surface, maneuverability and angle of rotation are important to avoid the craters and advance smoothly. Whereas for the case of other rough surfaces in Figure 10.F-H., speed is required to avoid the obstacles or move over them.

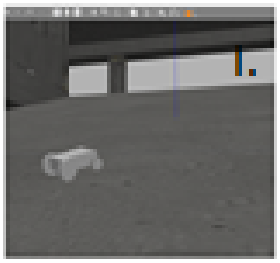


Figure 11: I. Slope

In the case of the plane slope in Figure 11.I, more speed should be provided than in the case of a plane surface as the robot needs to go to a higher place carrying its entire weight as the frictional force is acting in the opposite direction while moving up and in the direction of motion while coming down. But if the speed is given above a certain value the robot topples and if we rotate the body at high speed it topples.



Figure 12: J. Collapsed roof, K. Slippery Slope, L. Rough slope

In the case of the collapsed roof in Figure 12.J, and the rough slope in Figure 12.L, the body does not move easily as there is a frictional force between the surface and the robot compared to the case of plane slope and as a result, the body does not topple easily. But, for the case of driving terrain in Figure 12.K, the surface is damp and the robot slips and drifts while moving with more than the required speed but when given an appropriate speed the robot moves smoothly without sliding as the inclination of the plane is not constant.

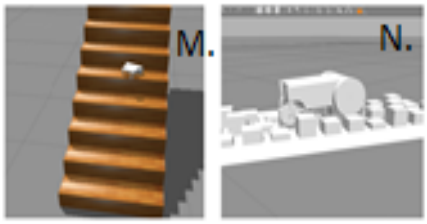


Figure 13: M. Steps, N. Uneven Path

Stairs are also a type of terrain but with both inclination and rough surface as shown in Figure 13.M, where control and mobility play an important part in the manoeuvrability of the robot. The uneven path in Figure 13.N is like rough paths in Figure 10.F-H, which requires more speed and mobility to avoid getting stuck in between the obstacles.

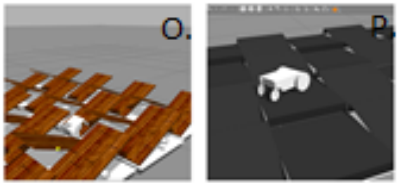


Figure 14: O. Teeter Ramps, P. Ramps

For testing the mobility and stability of the robot teeter ramps are chosen as shown in Figure 14.O The robot should move from one ramp to another while balancing the weight and move steadily over the ramps. Similarly, ramps were placed in Figure 15.P, to test the maneuverability of the robot to avoid obstacles and move smoothly over the ramps.

7. Results

The below Table provides the details of the frequencies at which the data is being published on different topics in different terrains and time taken for the robot to complete the first half of the distance as T1(sec) and the other half as T2(sec). From which we can infer that even in different scenarios the rate at which the data is being published is the same and the change in the behavior is mainly due to the speed and control, which affects the time taken by the robot to travel a certain distance.

Table 1: Frequencies of different Topics are published in multiple terrains

Test surface	Joy_ topic	Cmd_ vel	Joint_ states	Tf	Time (T1)	Time (T2)
A	24.42	24.86	9.97	54.87	10	10
B	24.14	24.94	9.98	55.0	11.8	11.8
C	24.33	24.92	9.92	55.04	12.4	12.4
D	24.01	24.66	9.96	54.99	10.7	10.7
E	24.23	24.53	10.0	54.88	11.2	11.2
F	24.46	24.79	9.98	55.09	13.6	13.6
G	24.39	24.35	9.96	54.89	14	14
H	24.19	24.95	9.99	54.93	19.6	19.5
I	24.53	24.92	10.0	54.91	12.2	7.93
J	24.36	24.61	9.97	54.99	14.7	11.9
K	24.29	24.42	9.99	54.89	13.2	9.1
L	24.12	24.84	10.0	55.02	15.6	12.3
M	24.24	24.18	9.92	54.9	-	10.8
N	24.2	24.4	10.0	55.01	16.2	17.5
O	24.18	24.49	9.76	54.96	15.3	15.7
P	24.09	24.26	10.0	54.88	17.4	16.9

While rotating the body the amount of speed required and the angle of rotation differs for different surfaces. When the body is moving in a straight line, the center point of rotation lies at an infinite distance from the robot. Similarly, when the angle of rotation of the robot increases the center point of rotation decreases until it approaches the center of the robot which works on the same principle of a differential drive plugin. When the angle of rotation for the robot is maximum the robot rotates at its own position as a center. When the angle of rotation increases the center point of rotation moves away from the robot and it occupies more space to rotate. As the outer wheel is at a larger distance from the center point of rotation the outer wheel covers more distance compared to the inner rotating wheel. This shows that angular orientation also plays a major role in the movement of the robot. Under normal circumstances, it might take a while to turn or control the robot in different conditions but using differential drive mechanism we can maneuver the robot easily through any type of terrains.

8. Conclusion

The major focus of this work is to model the differential drive for a wheeled robot. The mathematical equations have been derived for a robot with two caster wheels and two main drive wheels, by assuming that the casters are ideal and will not affect the drive in any form. The proposed model has been implemented using python and the file has been successfully converted to a ROS node for integrating it with ROS. The model has been evaluated using the simulation tools. A URDF file of the robot model has been generated from SolidWorks (SW) using SW to URDF exporter plugin and this has been triumphantly launched in the Gazebo platform and integrated with the node generated previously. The hardware integration of the steering wheel with the Gazebo environment has also been achieved. All the works referred to in this work are either dealing with the modeling of different aspects of different drive mechanisms or the ROS aspect of the robots, controllers, control systems. All the previous works referred to in this work discuss the usage of joysticks or have steering on the robot. As stated earlier, the major difference of this work is that the steering wheel and the robot are different entities. They do not have any physical contact. But virtually it is similar to having the steering physically on the robot. The ROS node takes care of the exchange of control information between these two entities. The robot has been tested in different environments and different readings from ROS have been collected to evaluate the performance, which shows a steady performance of the published velocity values and stable performance of the robot.

8.1. Future work

Gazebo simulation and RViz visualization with two-wheeled differential drive vehicle integrating with mapping modules and some sensors make the robot autonomous and increase the performance and efficiency of the robot. To implement and evaluate these applications on the real-time heavy-duty robot in the future plan. Improvising the performance of the robot by optimizing the algorithms which can reduce the delay in the response. Evaluating the performance of the heavy-duty may lead to further improvement of the robot. Optimizing the robot with high profile sensor components will get accurate and precise results.

Acknowledgment

Firstly, we would like to thank our institute Amrita Vishwa Vidyapeetham for providing the laboratory and components. The editors of HuTLabs and its staff have supported this work in each and every point of time and the people from the department of electronics and communication engineering have been constant supporters for the project at any point. Finally, we'd like to thank all the authors, maintainers and all the ROS users who worked and helped for the project with great positivity, composure, and determination.

References

[1] R. K. Megalingam, D. Nagalla, R. K. Pasumarthi, V. Gontu, and P. K. Allada, "ROS Based, Simulation and Control of a Wheeled Robot using Gamer's Steering Wheel, "2018 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2018 DOI:10.1109/CCAA.2018.8777569

[2] Wei Yu, Oscar Ylaya Chuy, Jr., Emmanuel G. Collins, Jr., Patrick Hollis, " Analysis and Experimental Verification for Dynamic Modeling of a Skid-Steered Wheeled Vehicle", IEEE Transactions on Robotics., vol. 26, pp. 340-353, Mar. 2010. DOI: 10.1109/TRO.2010.2042540

[3] Li Wang, Xinhui Liu, Xin Wang, Beibei Fu, Ran Xu, " Research on differential performance of four-wheel independent steering of a hydraulic wheel-driving off-road vehicle", The Journal of Engineering., vol. 2019, pp. 68-73, May. 2019. DOI:10.1049/joe.2018.8956

[4] Nobutake Hiraoka, Katsuhiko Inagaki, "A study of a new controller interface for omnidirectional robots", 10th Asian Control Conference (ASCC), 2015. DOI: 10.1109/ASCC.2015.7244592

[5] Xudong Ma Fang Fang; Kun Qian; Can Liang, "Networked robot systems for indoor service enhanced via ROS middleware", 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2018.DOI: 10.1109/ICIEA.2018.8397832

[6] Rajesh Kannan Megalingam; Sricharan Boddupalli; K. G. S. Apuroop, "Robotic arm control through mimicking of miniature robotic arm", 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), 2017. DOI:10.1109/ICACCS.2017.8014622

[7] Jianfeng Liao, Zheng Chen, Bin Yao, " Performance-Oriented Coordinated Adaptive Robust Control for Four-Wheel Independently Driven Skid Steer Mobile Robot", IEEE Access., vol. 5, pp. 19048-19057, Sept. 2017. DOI: 10.1109/ACCESS.2017.2754647

[8] Jianfeng Liao, Zheng Chen, Bin Yao, " Model-Based Coordinated Control of Four-Wheel Independently Driven Skid Steer Mobile Robot with Wheel-Ground Interaction and Wheel Dynamics.", IEEE Transactions on Industrial Informatics., vol. 15, pp. 1742-1752, Sept. 2018. DOI10.1109/TII.2018.2869573

[9] Masayoshi Wada, "A Joystick Steering Control System with Variable Sensitivity For Stable High Speed Driving", IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society, 2013. DOI: 10.1109/IECON.2013.6699781

[10] F.U. Rehman, M.M. Ahmed, " Steering Control Algorithm for a class of wheeled mobile robots", IET Control Theory and Applications., vol. 1, pp. 915-924, Jul. 2007. DOI: 10.1049/iet-cta:20060189

[11] M.A. Minor, B.W. Albiston, C.L. Schwensen, " Simplified Control of a two-axle compliant framed wheeled mobile robot", IEEE Transactions on Robotics, vol. 22, pp. 491-506, Jun. 2006. DOI: 10.1109/TRO.2006.875503

[12] Jingang Yi, Hongpeng Wang, Junjie Zhang, Dezhen Song, Suhada Jayasuriya, Jingtai Liu, " Kinematic Modeling and Analysis of Skid-Steered Mobile Robots with Applications to Low-Cost Inertial-Measurement -Unit-Based Motion Estimation", IEEE Transactions and Robotics., vol 25, pp. 1087-1097, Jul. 2009. DOI: 10.1109/ROBOT.2009.5152342

[13] Wei Shen, Zhichun Pan, Min Li, Hui Peng, " A Lateral control method for Wheeled-Footed Robot based on Sliding mode control and Steering Prediction", IEEE Access., vol. 6, pp. 58086-58095, Oct. 2018. DOI: 10.1109/ACCESS.2018.2873020

[14] Junnian Wang, Zheng Luo, Yan Wang, Bin Yang, Francis Assadian, " Coordinated Control of Differential Drive Assist Steering and Vehicle Stability Control for Four-Wheel-Independent-Drive EV", IEEE Transactions on Vehicular Technology., vol. 67, pp. 11453-11467, Oct. 2018. DOI 10.1109/TVT.2018.2872857

[15] Yousuf, L.S.Email Author," Experimental and simulation investigation of nonlinear dynamic behavior of a polydyne cam and roller follower mechanism", Mechanical Systems and Signal Processing 116, pp. 293-309, 2018. DOI: 10.1016/j.ymssp.2018.06.028

[16] Altarazi, S.A. Ammouri, M.M.," Concurrent manual-order-picking warehouse design: a simulation-based design of experiments approach", International Journal of Production Research pp. 1-19, 2017. DOI: 10.1080/00207543.2017.1421780

[17] Ashley L. Guinan; Nathaniel A. Caswell; Frank A. Drews, "William R. ProvancherA video game controller with skin stretch haptic feedback," 2013 IEEE International Conference on Consumer Electronics (ICCE), 2013. DOI: 10.1109/ICCE.2013.6486973

[18] Rajesh Kannan Megalingam; Sarath Sreekanth; Govardhan; Chinta Ravi Teja; Akhil Raj, "Wireless gesture controlled wheelchair", 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), 2017. DOI: 10.1109/ICACCS.2017.8014621

[19] Rajesh Kannan Megalingam; Deepak Nagalla; Pasumarthi Ravi Kiran; Ravi Teja Geesala; Katta Nigam, "Swarm based autonomous landmine detecting robots", 2017 International Conference on Inventive Computing and Informatics (ICICI), 2017. DOI: 10.1109/ICICI.2017.8365205

[20] Lentin Joseph, "learning Robotics using Python, Packt Publishing, 2015.

[21] W. Huang, X. Wu, Q. Zhang, N. Wu, and Z. Song, "Trajectory optimization of autonomous driving by differential dynamic programming," 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), Singapore, 2014 DOI: 10.1109/ICARCV.2014.7064582

[22] Y. Wu et al., "Experimental kinematics modeling estimation for wheeled skid-steering mobile robots," 2013 IEEE International Conference on Robotics and Biomimetics (ROBIO), Shenzhen, 2013 DOI: 10.1109/ROBIO.2013.6739470

[23] El-Dessoky, M.M, Alzahrani, E.O, Almohammadi, N.A, "Chaos control and function projective synchronization of novel chaotic dynamical system", Journal of Computational Analysis and Applications 27(1), pp. 162-172, 2018. DOI: 10.1155/2011/452671

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 280-286 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Electronic Steering of the Vehicle Axle with Force Feedback

Tomas Pawlenka*, Jaromir Skuta, Jiri Kulhanek

Faculty of Mechanical Engineering, Department of Control Systems and Instrumentation, VSB – Technical University of Ostrava, 708 00, Czechia

ARTICLE INFO

Article history:
Received: 14 January, 2020
Accepted: 29 February, 2020
Online: 20 March, 2020

Keywords:
steering wheel
force feedback
Steer-by-Wire
autonomous vehicles

ABSTRACT

The technology of electronic steering with force feedback is related to electronic vehicle control. This technology is also known as "steer-by-wire" technology and is specific in that there is not any mechanical connection between the axle and the steering wheel. This technology is currently not very widespread. However, developments in the automotive industry are slowly moving towards this technology in the context of autonomous vehicles. This study includes a review of current technologies in this area and suggests a technical solution for Steer-by-Wire technology and its verification in the real vehicle model in the university environment.

1. Introduction

The paper follows the work that was presented at the International Carpathian Control Conference (ICCC) [1]. This work dealt with the design of the electronic steering wheel and will be briefly described in this article as well. This article is further expanded by research about Steer-by-Wire technology and axle control using linear servos and a designed electronic control system. The main goal of this work is to make some progress in the area of Steer-by-Wire technology.

The technology of Steer-by-Wire is currently very widespread not only in the automotive industry. It is most often mentioned in connection with autonomous vehicles. It is a very special technology which must meet rules for safe and reliable operation. To ensure the safety of drivers and all road users, vehicles equipped with these systems must contain many electronics and sensors. However, the development of these technologies is still underway, and it will take some time to fully deploy to normal traffic.

In modern vehicles, systems that interfere with vehicle control are already implemented. However, we cannot call them autonomous. An example of such a system is a parking assistant, adaptive cruise control or a system that holds the vehicle in the traffic lane.

*Corresponding Author: Tomas Pawlenka, tomas.pawlenka@vsb.cz

Electronic vehicle control technology can be solved in many ways. It depends only on our requirements. As a first requirement, we can consider the independent steering or steering of both wheels simultaneously. Yamaguchi and Murakami [2] described a standard Steer-by-Wire model with one steering motor, which controls both wheels together via rack and pinion. It is based on a classic electric power steering system without a mechanical connection with a steering wheel. An example of a driving assistant system is described by Liu et al. [3] and it is based on the same principle. As another requirement can be considered the physical principle of the selected actuators. Ye et al. [4] explain the model of the vehicle's electro-hydraulic system that controls the steering of all four wheels.

Other issues in this area can provide feedback to the driver according to the road surface or advanced methods of control. The solution may be to implement force, torque and yaw rate sensors or accelerometers as described in [5-7]. A very interesting solution is also described by Fahami et al. [8]. They propose an algorithm to create force feedback torque based on the current measurement of the DC motor.

For advanced methods of control, a bilateral control scheme to improve vehicle handling can be mentioned, which was proposed by Zheng et. al. [9].

The main content of this paper is about the proposal of an electronic steering system technical solution, which is based on a

clear electric principle. This choice depends on a lower price than electrohydraulic systems. It utilizes independent steering of both front wheels with the use of power linear servos mounted on each wheel.

Figure 1 shows the block diagram of the chosen technology. It consists of a steering wheel unit, steering control system and two independent linear actuators. The steering wheel unit is designed for generating force feedback and setting the desired steering angle and the steering control system serves for controlling linear servos according to the desired steering angle set by a driver.

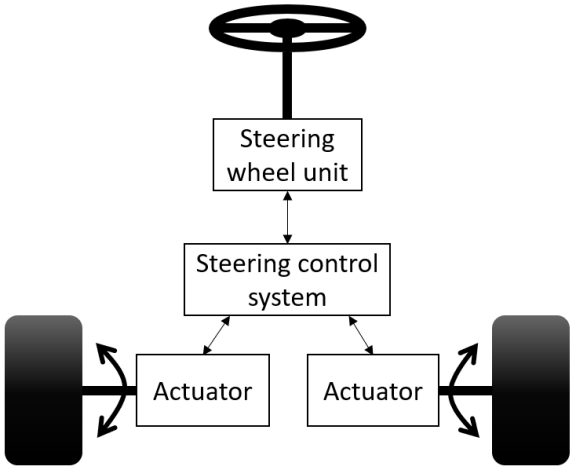


Figure 1 – Block structure of the proposed system

2. Design of steering wheel unit

In the beginning, the model of the steering wheel unit was designed. There are more options to create a steering wheel unit to control the steering and generate force feedback to the driver. In [2] the steering wheel rotation is proposed utilizing a motor and a belt drive. This method of torque transmission was also chosen in this work for easier implementation. The disadvantage of the gears is the play between the teeth, which in this case could cause steering insensitivity.

The main task of the proposed system is to provide the steering wheel position to the steering control system and generate force feedback to the driver. Force feedback is generated by a DC motor. The DC motor shaft is coupled to the steering wheel via a dual belt drive, which increases the final torque of the steering wheel.

The angle of the steering wheel is measured using the SEAT automotive position sensor. The advantage of this sensor is that it provides data over the CAN bus and it is also very reliable due to its robust design.

The Arduino UNO board was selected as the main control unit and it was interfaced with the CAN bus shield based on MCP2515 chip. The power of the DC motor is controlled by the dual H-bridge according to the PWM signal in the range of 0-5 V from the Arduino unit. Part of the design is also a current sensor. It measures the direction and value of the DC motor current. The current direction is important to determine if the driver affects the steering wheel with any force.

Figure 2 shows the whole structure of the steering wheel unit with the possibility of setting parameters through a computer application. As can be seen from the picture, the structure is divided into three levels according to their main task.

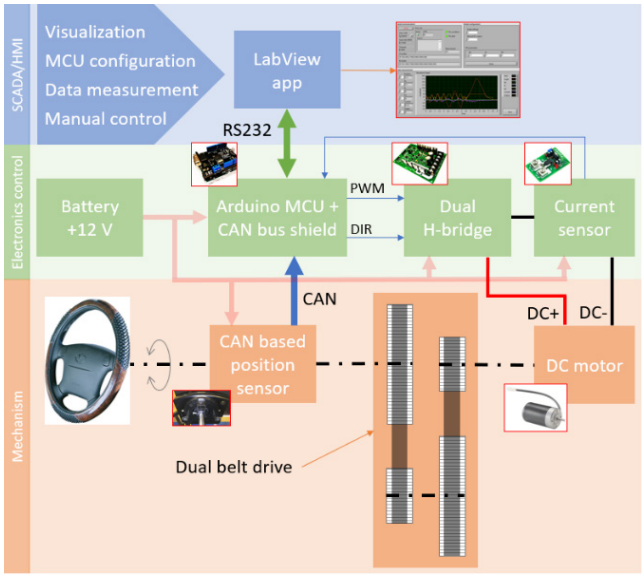


Figure 2 - Block diagram of the steering wheel unit

2.1. Construction design and assembly

As the feedback motor, the DOGA 162 was chosen. It is a motor that uses permanent magnets to generate a magnetic field. The principle is based on the mutual attraction of the positive and negative poles [10]. The initial torque of the chosen motor is about 1 Nm from the datasheet. Parameters like performance, size, weight, availability, and price played a major role in the choice of this DC motor.

As already mentioned, the power transfer of the feedback motor to the steering wheel shaft was designed by a dual belt drive. The belt drive was designed to meet the requirement for the resulting torque. First, the gear ratio was defined as the same for both gears as 4.5. HTD pulleys with the required number of teeth were then selected based on the gear ratio. The total gear ratio through both gears was calculated as 20.25. Then it was necessary to calculate the minimal axial distance between pulleys. The actual axial distance was then selected based on the standard lengths of the Continental HTD belts.

The equations and formulas for obtaining the belt length are described in [1].

The calculation of the resulting starting torque of the DC motor is also described in [1].

Finally, the initial torque from the total gear ratio has been calculated and it can be noted that its value meets the specified

requirement of 20 Nm. This value can be considered as enough for this application.

In Figure 3, the mechanical construction of the model is shown. The prototype frame is made of metal sheets and threaded rods. All shafts were mounted with the construction frame through ball bearings fixed in bearing holders. The position sensor was connected to the steering wheel shaft using a 3D printed connecting part.

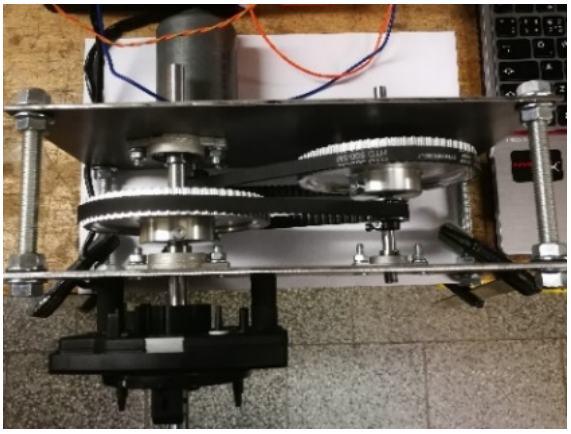
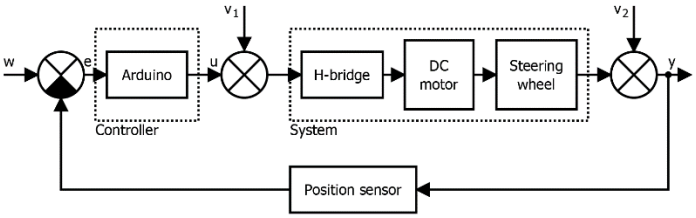


Figure 3 – Realized mechanism of the steering wheel unit

2.2. Electronics control system

The design of the electronic control system includes the selection of suitable components and their correct connection for safe operation.

Figure 4 describes a regulation loop showing which components represent the controller, the system and the measured feedback.



w - input value, e - control error, u - manipulated variable, y - processed variable
v_1 - fault value before the system, v_2 - fault value behind the system

Figure 4 – The regulation loop of the electronic steering wheel unit

The controller is the evaluation board Arduino UNO. The manipulated variable is generated by the PWM signal in the 0-5 V range. The system consists of a dual H-bridge, DC motor, dual belt drive, and steering wheel. The dual H-bridge controls the power of the DC motor based on the manipulated variable. The steering wheel position reference value can be considered as an input. The actual controlled position can be considered as the output. The position feedback is measured by the already mentioned sensor based on CAN.

The CAN bus is a communication standard used for data transfer in the automotive industry. A twisted-pair copper line is usually used as the transmission medium [12].

The position sensor provides data about the current position and speed of rotation. The sensor resolution is 240 values per revolution. Measurement of its ranges was performed using the PCAN-USB interface from Peak System.

The Arduino unit was developed as an open-source platform. For programming, it is possible to use the development environment Arduino, which is free. The main function of this unit is to measure and control the position of the steering wheel. For this purpose, the PID control algorithm was implemented in the Arduino. Another function is cyclical data sending via a serial interface to a computer application. In addition to sending data, the unit can also receive data in the form of configuration commands from the application.

The mentioned application was developed in LabView from National Instruments. This environment uses a block structure for programming. The final design is shown in Figure 5.

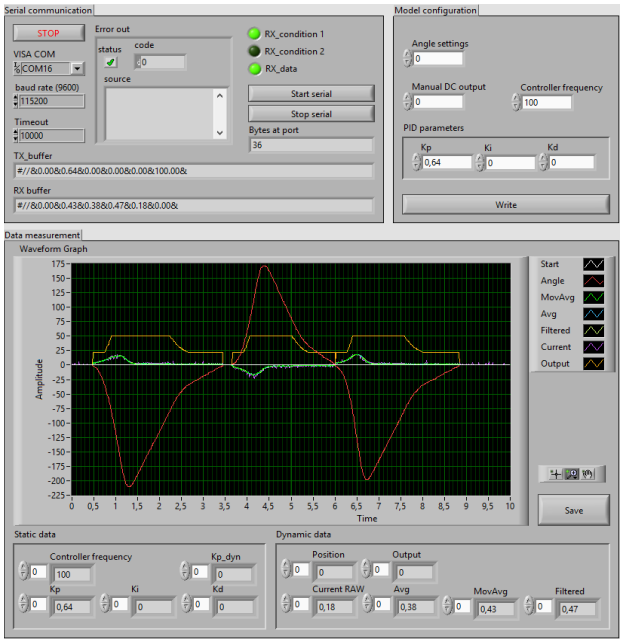


Figure 5 – Steering wheel unit application

The application was designed and divided into three sections. The first section consists of serial communication settings, error indicators and the display of received and sent buffers to or from the Arduino unit.

The second section is for model configuration and manual control. There is the possibility of manually adjusting the steering wheel angle as an offset position or DC motor power in the 0-255 PWM range to measure step response. This section also contains a form for setting PID controller parameters and controller frequency.

The last section includes tools for data measurement and visualization. Except for the graph, which shows the last values on the selected timeline, this section also contains numeric indicators of current values. Measured data can be saved to TDMS file using the save button.

2.3. System identification and controller design

An important part of the work is to determine the parameters of the proposed model. Deterministic methods based on measured step response were chosen for system identification. For step response measurement, the developed application was used. That was measured during the steering wheel rotation of 360 degrees. The step size was set to match the maximum motor power. According to the step response, it was found out that it is the integration system. It will not stabilize.

The inertia integration system can be described by the following form of the transfer function [13]:

$$G(s) = \frac{1}{T_I s} \cdot \frac{1}{(T_1 s + 1)^n} \tag{1}$$

Integral time and inertia were obtained using standard identification methods. The calculations revealed that it is a first-order system defined by the following form of the transfer function [13]:

$$G(s) = \frac{k_1}{s \cdot (T_1 s + 1)} \tag{2}$$

Subsequently, a step response was simulated using a calculated transfer function in Matlab Simulink. The accuracy of the simulated step response was verified by comparison with the measured step response. Figure 6 shows that the transfer function parameters have been calculated correctly.

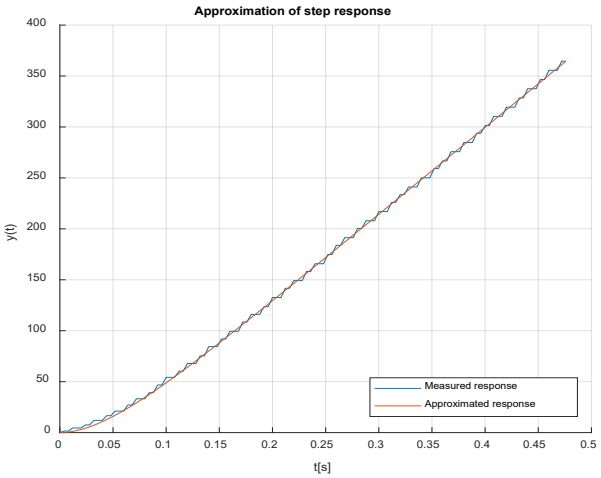


Figure 6 – Step response approximation

The controller parameters were then calculated for the derived transfer function. The Desired model method was used for this purpose. This method established a PD controller for the first-order integration system.

The following equation shows the transfer function of a conventional PD controller in the discrete form [14]:

$$G_R(z) = k_p \cdot \left(1 + \frac{T_D}{T} \cdot \frac{z-1}{z}\right) \tag{3}$$

Where:

k_p – proportional gain,
T_D – derivative time
T – controller sample time

The controller parameters were obtained for 0 % overshoot and 100 Hz sampling frequency. The calculation of the parameters is given by the following formulas [14]:

$$T_D^* = T_1 - \frac{T}{2}$$
$$k_p^* = \frac{1}{\alpha \cdot T \cdot k_1} \tag{4}$$

Where:

T_1 – transfer function time constant,
k_1 – transfer function coefficient,
T – sample rate
α – table coefficient

Figure 7 shows the process of control with proposed parameters while turning the steering wheel. To facilitate steering wheel rotation, the manipulated variable was limited to 20% of DC motor power.

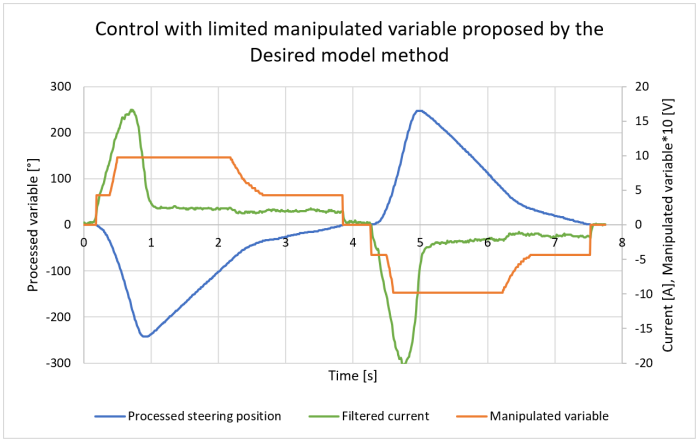


Figure 7 – Final process of control

The graph also contains the measurement of feedback motor load using the current sensor based on ACS758. Due to high noise, the signal was adjusted using an exponential filter.

The exponential filter is described by equation below [15]:

$$y_F(i) = K_F \cdot y_M(i) + (1 - K_F) \cdot y_F(i-1)$$
$$y_F(1) = y_M(i), 0 < K_F \leq 1 \tag{5}$$

Where:

$y_F(i)$ – new filtered value,
K_F – filtration coefficient,
$y_M(i)$ – measured value,
$y_F(i\text{-}1)$ – previous filtered value

The value of the filtration coefficient was chosen as 0.03.

Figure 8 shows the different methods of current signal conditioning. There is a comparison of moving average and exponential filtration.

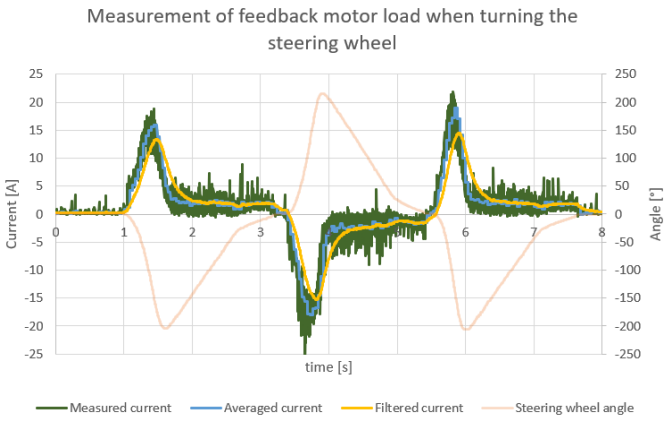


Figure 8 - Feedback motor load while turning the steering wheel

The solution for the future implementation of other systems such as parking assistance or vehicle stabilization can be designed advanced control with dynamic change of controller parameters depending on the required power of the motor.

In Figure 9, the prototype model of the electronic steering wheel is shown.

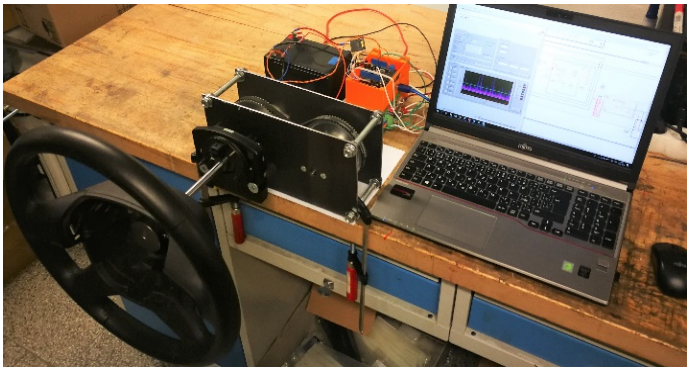


Figure 9 – Electronic steering wheel unit prototype model

3. Design of the steering control system

In the design of the steering control system, it was necessary to select appropriate components such as actuators or control unit. As the actuators, the super duty linear servos were selected for the wheel steering. These actuators were selected according to the calculated parameters described in [16]. This linear actuator can exert a force greater than 2 kN and includes a 10 kΩ potentiometer to provide information about the current position. For the control of linear drives, the cRIO from National Instruments together with Roboclaw motor driver was chosen.

The same procedure as for the steering wheel unit was performed to determine the PID controller parameters. This means that experimental identification was performed first. Then the controller parameters were derived for the determined transfer function [16].

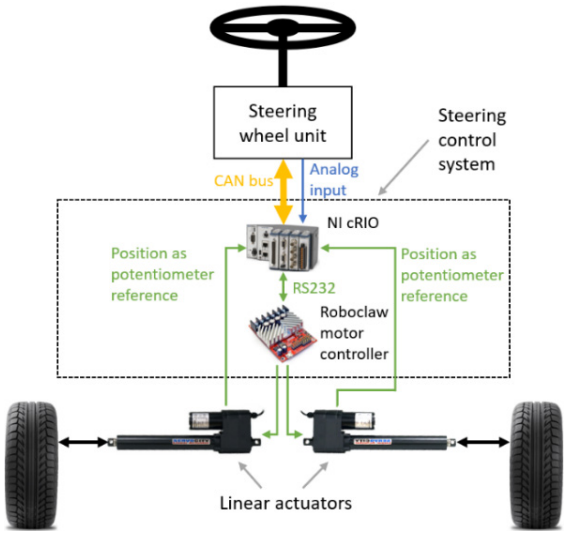


Figure 10 – Structure of the proposed electronic steering system, [16]

Figure 10 shows the whole structure of the proposed steering control system with selected components.

Figure 11 shows the linear actuators mounted on the axles of the test vehicle.



Figure 11 – Picture of testing vehicle with steer-by-wire technology

Finally, the program of the control system was created in LabView. The cRIO was not only used for testing, but also as a main unit of the steering control system. The control system also implements the condition of Ackermann steering geometry. The calculation of the Ackermann formula is described in [16].

Ackermann steering geometry is based on the different steering angles of the inner and outer wheels when the vehicle turns. The aim is that the turn radius of both wheels ideally has a

common midpoint. As a result, the tires are not so stressed when cornering the vehicle. The principle and more detailed information are described by Zhao et al. [17]. They deal with the design of the steering mechanism based on Ackermann.

4. Results and future work

This chapter summarizes the results of this work During this work, two systems were developed which communicate with each other. The first is the steering wheel unit, which is used to set the desired position of wheels and generate force feedback back to the driver. The maximum torque of the feedback DC motor was calculated as 20 Nm. The second system controls the linear actuators of both wheels.

Figure 12 shows the final relation between the steering angle of both wheels and the steering wheel position. The difference between wheels is given by Ackermann steering geometry of our car.

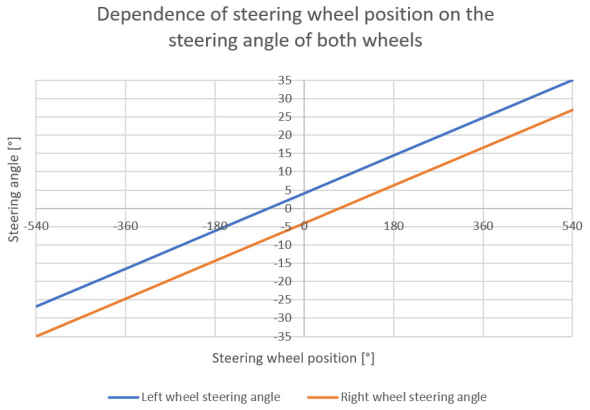


Figure 12 – Relationship between steering angle and steering wheel position

The goal of further development is to find a solution for road surface feedback and get it to the driver through the steering wheel. The solution can be force or torque sensors in combination with accelerometers or yaw rate sensors mounted on the wheels of the vehicle as described in [18, 19].

5. Conclusion

The current automotive world is moving towards autonomy and the replacement of mechanical systems by electronic ones. Therefore, this work was focused on the design and implementation of electronic vehicle control without a mechanical connection between the axle and the steering wheel. In this work were developed and implemented two systems that can communicate with each other. The first system is the steering wheel unit and the second is the steering control system. The steering wheel unit represents a steering wheel coupled to a DC motor using a dual belt drive and the position is measured by CAN based sensor. The main role of this system is to transmit the requirement for the steering angle of the wheels in the form of an analog signal to the steering control system. The other role is to generate force feedback back to the driver by the steering wheel. The steering control system is represented by the control unit, motor controller and two linear actuators for turning the wheels.

The result of this work is the possibility of independent control of both wheels according to the steering wheel position. The advantage of independent control is software geometry adjustment. The proposed system was tested on the real model of the electric car at the university.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by the European Regional Development Fund in the Research Centre of Advanced Mechatronic Systems project, CZ.02.1.01/0.0/0.0/16_019/0000867 within the Operational Programme Research, Development and Education and the project SP2020/57 Research and Development of Advanced Methods in the Area of Machines and Process Control supported by the Ministry of Education, Youth and Sports.

References

[1] T. Pawlenka and J. Škuta, "Electronic Steering With Force Feedback," 2019 20th International Carpathian Control Conference (ICCC), Krakow-Wieliczka, Poland, 2019, pp. 1-5. doi: 10.1109/CarpathianCC.2019.8765925

[2] Y. Yamaguchi and T. Murakami, "Adaptive Control for Virtual Steering Characteristics on Electric Vehicle Using Steer-by-Wire System", IEEE Transactions on Industrial Electronics, vol. 56, no. 5, pp. 1585-1594, 2009. Available: 10.1109/tie.2008.2010171.

[3] X. Liu et al., "Driving Assist System for Ultra-Compact EVs—Fundamental Consideration of Muscle Burden Owing to Differences in the Drivers' Physiques", Actuators, vol. 7, no. 3, p. 44, 2018. Available: 10.3390/act7030044.

[4] M. Ye, Q. Wang and S. Jiao, "RobustH2/H∞Control for the Electrohydraulic Steering System of a Four-Wheel Vehicle", Mathematical Problems in Engineering, vol. 2014, pp. 1-12, 2014. Available: 10.1155/2014/208019.

[5] G. Yin, N. Chen and P. Li, "Improving Handling Stability Performance of Four-Wheel Steering Vehicle via μ-Synthesis Robust Control", IEEE Transactions on Vehicular Technology, vol. 56, no. 5, pp. 2432-2439, 2007. Available: 10.1109/tvt.2007.899941.

[6] H. Yan, W. Zhang and D. Wang, "Wheel Force Sensor-Based Techniques for Wear Detection and Analysis of a Special Road", Sensors, vol. 18, no. 8, p. 2493, 2018. Available: 10.3390/s18082493.

[7] J. Pytka et al., "Determining Wheel Forces and Moments on Aircraft Landing Gear with a Dynamometer Sensor", Sensors, vol. 20, no. 1, p. 227, 2019. Available: 10.3390/s20010227.

[8] S. Fahami, H. Zamzuri and S. Mazlan, "Development of Estimation Force Feedback Torque Control Algorithm for Driver Steering Feel in Vehicle Steer by Wire System: Hardware in the Loop", International Journal of Vehicular Technology, vol. 2015, pp. 1-17, 2015. Available: 10.1155/2015/314597.

[9] H. Zheng, J. Hu and Y. Liu, "A bilateral control scheme for vehicle steer-by-wire system with road feel and steering controller design", Transactions of the Institute of Measurement and Control, vol. 41, no. 3, pp. 593-604, 2017. Available: 10.1177/0142331217734502.

[10] K. Kozeka, "A Motor Driven by Permanent Magnets Alone; A Clean and Abundant Source of Electromagnetic Energy from Iron and Other Ferromagnetic Materials", Natural Science, vol. 09, no. 09, pp. 319-329, 2017. Available: 10.4236/ns.2017.99031.

[11] V. Bhandari, Design of machine elements. New Delhi: McGraw-Hill Education (India), 2017.

[12] J. Škuta and J. Kulhánek, "Controll of car LED lights by CAN/LIN bus," Proceedings of the 2015 16th International Carpathian Control Conference (ICCC), Szilvasvarad, 2015, pp. 486-489. doi: 10.1109/CarpathianCC.2015.7145128.

[13] P. Noskievič, Modelování a identifikace systémů. Ostrava: Montanex, 1999.

[14] M. Vítečková, Základy automatické regulace, Přeprac. 2. vyd. Ostrava: VŠB - Technická univerzita Ostrava, 2008.

[15] H. Ravinder, "Forecasting With Exponential Smoothing Whats The Right Smoothing Constant?", Review of Business Information Systems (RBIS), vol. 17, no. 3, pp. 117-126, 2013. Available: 10.19030/rbis.v17i3.8001.
[16] Lawrance, “Steering by Wire for Sport Car”, Diploma thesis, Ostrava, 2019.
[17] J. Zhao, X. Liu, Z. Feng and J. Dai, "Design of an Ackermann-type steering mechanism", Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, vol. 227, no. 11, pp. 2549-2562, 2013. Available: 10.1177/0954406213475980.
[18] J. Iqbal, K. Zuhaib, C. Han, A. Khan and M. Ali, "Adaptive Global Fast Sliding Mode Control for Steer-by-Wire System Road Vehicles", Applied Sciences, vol. 7, no. 7, p. 738, 2017. Available: 10.3390/app7070738.
[19] J. Tian, J. Tong and S. Luo, "Differential Steering Control of Four-Wheel Independent-Drive Electric Vehicles", Energies, vol. 11, no. 11, p. 2892, 2018. Available: 10.3390/en11112892.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 284-290 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Controller Design Using Backstepping Algorithm for Fixed-Wing UAV with Thrust Vectoring System

Shogo Hirano*, Kenji Uchiyama, Kai Masuda

Nihon University, Department of Aerospace Engineering, Chiba 274-8501, Japan

ARTICLE INFO

Article history:
Received: 15 January, 2020
Accepted: 02 May, 2020
Online: 29 May, 2020

Keywords:
Backstepping control
Thrust vectoring system
Fixed-wing UAV

ABSTRACT

This paper describes the design method of a nonlinear flight controller for a fixed-wing UAV with a thrust vectoring system (TVS) using the backstepping method. The flight dynamics of the UAV exhibits strong nonlinear coupling behavior between its translational and rotational motion. The backstepping algorithm has been successfully applied to controller design for such a nonlinear system. However, the main idea of the method is to use some of the state variables as virtual control inputs that need ungeneratable forces by the UAV. To overcome this problem, we use the TVS that can generate thrust in an arbitrary direction. Numerical simulation is performed to confirm the effectiveness of the proposed control method for a fixed-wing UAV with the TVS.

1. Introduction

Unmanned aerial vehicles (UAVs) have been used for national projects, and multirotor are already introduced in various scenes. The fixed-wing UAVs have been used in a specific field [1]. The UAV has some advantages such as high cruising ability or large capacity of payloads. In particular, a small UAV with a wingspan of several meters is superior to a UAV with a wingspan of dozens of meters in maneuverability and maintainability. However, a small UAV has the problem that it is easily affected by the wind disturbance during a flight because of its small moment of inertia. As a result, it is difficult to achieve stable flight of the small UAV when a strong wind blows in a real mission. However, the small UAV cannot avoid an unexpected situation such as a strong wind during a flight because a high-performance device that measures the flight environment cannot be used due to its payload capacity.

One of the ways to solve the problem is to apply a nonlinear control method. The backstepping method [2]–[3], the nonlinear control methods, has also been studied in the fixed-wing UAV flying under the wind disturbance [4]–[7].The advantage of this method is that it makes separated subsystems stabilized by putting the input successively. Therefore, in this study, we attempt to control translational and rotational motion individually to implement the flight at the constant attitude angle under wind disturbance. However, there is a problem that the conventional fixed-wing UAV has a structural restriction that the altitude and the pitch angle cannot be controlled at the same time.

Therefore, we introduce the thrust vectoring system (TVS) for the UAV to solve the problem. The TVS has been used to operate the engine nozzle on aircraft that require maneuverability such as fighter aircraft but also a study on the UAV mounting a propeller-type TVS has been done at late years [8]–[13]. However, in those studies, rather than getting maneuverability as a fighter aircraft, the TVS is mainly used to control the attitude of a tail-sitter UAV during transition flight or the fixed-wing UAV with a broken dynamic surface. In this research, we apply the TVS to the generation of translational force to fly stably under wind disturbance. Finally, we confirm the effectiveness of the proposed method by the numerical simulation.

2. Nonlinear model of UAV

The nonlinear model of UAV can be described separately in translational motion and rotational motion. Control inputs are the thrust vector and the steering angles of an aileron, a rudder, and an elevator. The body coordinate and the inertial coordinate system of UAV are defined as shown in Fig.1.

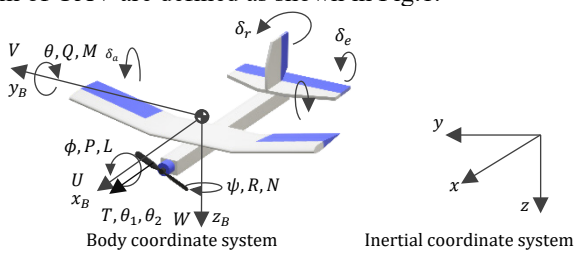


Figure 1: Definition of the coordinate system

2.1. Translational motion

The navigation equation of the UAV is given by

* Shogo Hirano, 7-24-1 Narashinodai, Funabashi, Chiba 274-8501, +81-47-469-5448, cssh19023@g.nihon-u.ac.jp

https://dx.doi.org/10.25046/aj050337

$$\dot{\mathbf{x}}_1 = \mathbf{g}_1(\mathbf{x}_3)\mathbf{x}_2 \tag{1}$$

where $\mathbf{x}_1 = [x\ y\ z]^{\mathrm{T}}$, $\mathbf{x}_2 = [U\ V\ W]^{\mathrm{T}}$, and $\mathbf{x}_3 = [\phi\ \theta\ \psi]^{\mathrm{T}}$ are the position of the UAV in the inertial coordinate system, its velocity in the body coordinate system, and the Euler angle, respectively. $\mathbf{g}_1(\mathbf{x}_3)$ is defined as

$$\boldsymbol{g}_1(\boldsymbol{x}_3) = \begin{bmatrix} c(\psi)c(\theta) & c(\psi)s(\theta)s(\phi) - s(\psi)c(\phi) & c(\psi)s(\theta)c(\phi) + s(\psi)s(\phi) \\ s(\psi)c(\theta) & s(\psi)s(\theta)s(\phi) + c(\psi)c(\phi) & s(\psi)s(\theta)c(\phi) - c(\psi)s(\phi) \\ -s(\theta) & c(\theta)s(\phi) & c(\theta)c(\phi) \end{bmatrix}$$

Here, c(·) and s(·) are abbreviations for the function cos(·) and the function sin(·), respectively.

Next, the translational motion equation of UAV is expressed by

$$\dot{\mathbf{x}}_2 = \mathbf{f}_2(\mathbf{x}_4)\mathbf{x}_2 + \mathbf{g}_2(\mathbf{x}_3) + \mathbf{F}_a + \mathbf{F}_u \tag{2}$$

where $\mathbf{x}_4 = [P\ Q\ R]^{\mathrm{T}}$ is the angular velocity of the body coordinate system, and nonlinear matrices $\mathbf{f}_2(\boldsymbol{x}_4)$, nonlinear vector $\mathbf{g}_2(\boldsymbol{x}_3)$ are defined as

$$\mathbf{f}_2(\mathbf{x}_4) = \begin{bmatrix} 0 & R & -Q \\ -R & 0 & P \\ Q & -P & 0 \end{bmatrix},\ \mathbf{g}_2(\mathbf{x}_3) = g\begin{bmatrix} -s\theta \\ s\phi c\theta \\ c\phi c\theta \end{bmatrix}$$

$\mathbf{F}_a = [X_a\ Y_a\ Z_a]^{\mathrm{T}}/m$ is the aerodynamic force vector, $\mathbf{F}_u = [X_u\ Y_u\ Z_u]^{\mathrm{T}}/m$ is the control input vector, and m is the mass of the UAV. Each element of $\mathbf{F}_a$ and $\mathbf{F}_u$ is the force along each axis.

2.2. Rotational motion

The following equation expressed the kinematics equation of the UAV.

$$\dot{\mathbf{x}}_3 = \mathbf{g}_3(\mathbf{x}_3)\mathbf{x}_4 \tag{3}$$

Nonlinear matrix $\mathbf{g}_3(\mathbf{x}_3)$ is defined as

$$\mathbf{g}_3(\mathbf{x}_3) = \begin{bmatrix} 1 & s\phi t\theta & c\phi t\theta \\ 0 & c\phi & -s\phi \\ 0 & s\phi/c\theta & c\phi/c\theta \end{bmatrix}$$

where t(·) denotes the function tan(·).

The rotational motion equation of the UAV is given as

$$\dot{\mathbf{x}}_4 = \mathbf{f}_4(\mathbf{x}_4)\mathbf{x}_4 + \mathbf{u}_a + \mathbf{u}_{st} \tag{4}$$

Nonlinear matrix $\mathbf{f}_4(\mathbf{x}_4)$ is defined as

$$\mathbf{f}_4(\mathbf{x}_4) = \mathbf{J}^{-1}\begin{bmatrix} 0 & R & -Q \\ -R & 0 & P \\ Q & -P & 0 \end{bmatrix}\mathbf{J}$$

$\mathbf{J}$ is the moment of inertia. $\mathbf{u}_a = \mathbf{J}^{-1}[L_a\ M_a\ N_a]^{\mathrm{T}}$ is the aerodynamic moment vector, and $\mathbf{u}_{st} = \mathbf{J}^{-1}[L_u\ M_u\ N_u]^{\mathrm{T}}$ is the control input vector. Each element of $\mathbf{u}_a$ and $\mathbf{u}_{st}$ is the moment of the force around each axis.

3. Controller design

In the backstepping method, the control system is designed considering the stability of each subsystem. In this paper, a control system is designed by putting (1)-(4) as subsystems S1 to S4 ,and its construction is as shown in Fig.2.

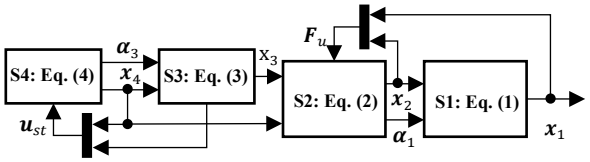


Figure 2: The structure of subsystem

In the above figure, control systems are constructed with two extensive subsystems that consist of the translational subsystems (S1, S2) and the rotational subsystems (S3, S4).

3.1. Translational control system

First, in order to give a target value $\mathbf{x}_{1c}$, the position, and velocity for UAV, we define error equation $\mathbf{x}_{1e}$ as follows:

$$\mathbf{x}_{1e} = \mathbf{x}_1 - \mathbf{x}_{1c} \tag{5}$$

When giving a target value of the velocity, the target value of the position that is the integrated value of the velocity is given instead, assuming the target value of velocity as steady. Then, to introduce a candidate of control Lyapunov function for the subsystem S1, V_1 is expressed as

$$V_1(\mathbf{x}_{1e}) = \frac{1}{2}\mathbf{x}_{1e}^{\mathrm{T}}\mathbf{x}_{1e} \tag{6}$$

The derivative of V_1 is

$$\begin{aligned} \dot{V}_1(\mathbf{x}_{1e}) &= \mathbf{x}_{1e}^{\mathrm{T}}\dot{\mathbf{x}}_{1e} \\ &= \mathbf{x}_{1e}^{\mathrm{T}}(\dot{\mathbf{x}}_1 - \dot{\mathbf{x}}_{1c}) \\ &= \mathbf{x}_{1e}^{\mathrm{T}}(\mathbf{g}_1(\mathbf{x}_3)\mathbf{x}_2 - \dot{\mathbf{x}}_{1c}) \end{aligned} \tag{7}$$

Here, considering state variable $\mathbf{x}_2$ as virtual input of S1, and virtual input is defined as $\boldsymbol{\alpha}_1(\mathbf{x}_1, \mathbf{x}_3, t)$. When $\dot{V}_1 < 0$, the Lyapunov stability theory is satisfied, and V_1 becomes a Lyapunov function. i.e., the subsystem S1 gets stable. The following equation can determine the virtual input $\boldsymbol{\alpha}_1(\mathbf{x}_1, \mathbf{x}_3, t)$ that satisfies the condition for the system stability because $\mathbf{g}_1(\mathbf{x}_3)$ is a regular matrix.

$$\boldsymbol{\alpha}_1(\mathbf{x}_{1e}, \mathbf{x}_3, t) = \mathbf{g}_1(\mathbf{x}_3)^{-1}\{\dot{\mathbf{x}}_{1c} - \mathbf{K}_1\mathbf{x}_{1e}\} \tag{8}$$

where $\mathbf{K}_1$ is a positive definite and diagonal matrix.

The error equation between state variable $\mathbf{x}_2$ and virtual input $\boldsymbol{\alpha}_1$ is as (9).

$$\mathbf{z}_1 = \mathbf{x}_2 - \boldsymbol{\alpha}_1 \tag{9}$$

Then, we define V_2 to introduce a candidate as a control Lyapunov function for the whole of the translational systems.

$$V_2(\mathbf{x}_{1e}, \mathbf{z}_1) = V_1 + \frac{1}{2}\mathbf{z}_1{}^{\mathrm{T}}\mathbf{z}_1 \tag{10}$$

The derivative of V_2 is

$$\begin{aligned}\dot{V}_2(\mathbf{x}_{1e},\mathbf{z}_1) &= \mathbf{x}_{1e}^{\mathrm{T}}\{\mathbf{g}_1(\mathbf{x}_3)(\mathbf{z}_1+\boldsymbol{\alpha}_1)-\dot{\mathbf{x}}_{1c}\}+\mathbf{z}_1{}^{\mathrm{T}}\dot{\mathbf{z}}_1\\ &= \mathbf{x}_{1e}^{\mathrm{T}}(\mathbf{g}_1(\mathbf{x}_3)\boldsymbol{\alpha}_1-\dot{\mathbf{x}}_{1c})\\ &\quad+\mathbf{z}_1{}^{\mathrm{T}}\left\{\dot{\mathbf{z}}_1+\left(\mathbf{g}_1(\mathbf{x}_3)\right)^{\mathrm{T}}\mathbf{x}_{1e}\right\}\\ &= \mathbf{x}_{1e}^{\mathrm{T}}(\mathbf{g}_1(\mathbf{x}_3)\boldsymbol{\alpha}_1-\dot{\mathbf{x}}_{1c})\\ &\quad+\mathbf{z}_1{}^{\mathrm{T}}\left\{\mathbf{f}_2(\mathbf{x}_4)\mathbf{x}_2+\mathbf{g}_2(\mathbf{x}_3)+\mathbf{F}_a+\mathbf{F}_u\right.\\ &\quad\left.-\dot{\boldsymbol{\alpha}}_1+\left(\mathbf{g}_1(\mathbf{x}_3)\right)^{\mathrm{T}}\mathbf{x}_{1e}\right\}\end{aligned} \tag{11}$$

The first term on the right side of (11) is negative-definite according to (8). Therefore, putting the input $\mathbf{F}_{uc}$ as written in (12), the second line of the summation of (11) becomes negative-definite that $\dot{V}_2 < 0$ is satisfied, and V_2 is determined as Lyapunov function.

$$\begin{aligned}\mathbf{F}_{uc} = -\mathbf{K}_2\mathbf{z}_1-\mathbf{f}_2(\mathbf{x}_4)\mathbf{x}_2-\mathbf{g}_2(\mathbf{x}_3)-\mathbf{F}_a+\dot{\boldsymbol{\alpha}}_1\\ -\left(\mathbf{g}_1(\mathbf{x}_3)\right)^{\mathrm{T}}\mathbf{x}_{1e}\end{aligned} \tag{12}$$

where $\mathbf{K}_2$ is a positive definite and diagonal matrix.

3.2. Rotational control system

Similar to the previous section, the error equation $\mathbf{x}_{3e}$ is defined by using the target value $\mathbf{x}_{3c}$ of the Euler angle.

$$\mathbf{x}_{3e} = \mathbf{x}_3 - \mathbf{x}_{3c} \tag{13}$$

The candidate of control Lyapunov function, V_3 for S3 is defined as

$$V_3(\mathbf{x}_{3e}) = \frac{1}{2}\mathbf{x}_{3e}^{\mathrm{T}}\mathbf{x}_{3e} \tag{14}$$

The time derivative of V_3 is

$$\begin{aligned}\dot{V}_3(\mathbf{x}_{3e}) &= \mathbf{x}_{1e}^{\mathrm{T}}\dot{\mathbf{x}}_{1e} = \mathbf{x}_{3e}^{\mathrm{T}}(\dot{\mathbf{x}}_3-\dot{\mathbf{x}}_{3c})\\ &= \mathbf{x}_{3\mathrm{e}}^{\mathrm{T}}(\mathbf{g}_3(\mathbf{x}_3)\mathbf{x}_4-\dot{\mathbf{x}}_{3c})\end{aligned} \tag{15}$$

In (15), considering the state variable $\mathbf{x}_4$ as a virtual input of S3, we replace it by a virtual input $\boldsymbol{\alpha}_3(\mathbf{x}_3,\mathbf{x}_4,t)$. The following equation gives a virtual input that satisfies the Lyapunov stability theory.

$$\boldsymbol{\alpha}_3(\mathbf{x}_{3e},\mathbf{x}_4,t) = \mathbf{g}_3(\mathbf{x}_3)^{-1}(\dot{\mathbf{x}}_{3c}-\mathbf{K}_3\mathbf{x}_{3e}) \tag{16}$$

where $\mathbf{K}_3$ is a positive definite and diagonal matrix.

The error equation between state variable $\mathbf{x}_4$ and virtual input $\boldsymbol{\alpha}_3$ is defined as

$$\mathbf{z}_3 = \mathbf{x}_4 - \boldsymbol{\alpha}_3 \tag{17}$$

Then, the candidate of control Lyapunov function for the rotational motion is defined as

$$V_4(\mathbf{x}_{3e},\mathbf{z}_3) = V_3 + \frac{1}{2}\mathbf{z}_3{}^{\mathrm{T}}\mathbf{z}_3 \tag{18}$$

The time derivative of V_4 is obtained as follows:

$$\begin{aligned}\dot{V}_4(\mathbf{x}_{3e},\mathbf{z}_3) &= \mathbf{x}_{3e}^{\mathrm{T}}\{\mathbf{g}_3(\mathbf{x}_3)(\mathbf{z}_3+\boldsymbol{\alpha}_3)-\dot{\mathbf{x}}_{3c}\}+\mathbf{z}_3{}^{\mathrm{T}}\dot{\mathbf{z}}_3\\ &= \mathbf{x}_{3e}^{\mathrm{T}}(\mathbf{g}_3(\mathbf{x}_3)\boldsymbol{\alpha}_3-\dot{\mathbf{x}}_{3c})\\ &\quad+\mathbf{z}_3{}^{\mathrm{T}}\left\{\dot{\mathbf{z}}_3+\left(\mathbf{g}_3(\mathbf{x}_3)\right)^{\mathrm{T}}\mathbf{x}_{3e}\right\}\\ &= \mathbf{x}_{3e}^{\mathrm{T}}(\mathbf{g}_3(\mathbf{x}_3)\boldsymbol{\alpha}_3-\dot{\mathbf{x}}_{3c})\\ &\quad+\mathbf{z}_3{}^{\mathrm{T}}\left\{\mathbf{f}_4(\mathbf{x}_4)\mathbf{x}_4+\mathbf{u}_a+\mathbf{u}_{st}-\dot{\boldsymbol{\alpha}}_3\right.\\ &\quad\left.+\left(\mathbf{g}_3(\mathbf{x}_3)\right)^{\mathrm{T}}\mathbf{x}_{3e}\right\}\end{aligned} \tag{19}$$

By calculating the input command value $\mathbf{u}_{stc}$ to be as the following equation, the second and third lines of the last summation of (19) becomes negative-definite, and it makes the rotational system stable.

$$\mathbf{u}_{stc} = -\mathbf{K}_4\mathbf{z}_3-\mathbf{f}_4(\mathbf{x}_4)\mathbf{x}_4-\mathbf{u}_a+\dot{\boldsymbol{\alpha}}_3-\left(\mathbf{g}_3(\mathbf{x}_3)\right)^{\mathrm{T}}\mathbf{x}_{3e} \tag{20}$$

where $\mathbf{K}_4$ is a positive definite and diagonal matrix.

Fig.3 Shows the block diagram that includes the controller obtained in sections 3.1 and 3.2.

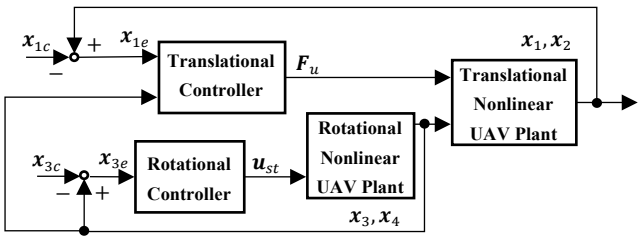


Figure 3: Block diagram of the proposed controller

3.3. Stability analysis

In this section, we analyze the stability of the entire control system that we designed. The candidate of control Lyapunov function of the entire control system is defined.

$$\begin{aligned}V &= V_1+V_2+V_3+V_4\\ &= 2V_1+\frac{1}{2}\mathbf{z_1}^{\mathrm{T}}\mathbf{z_1}+2V_3+\frac{1}{2}\mathbf{z_3}^{\mathrm{T}}\mathbf{z_3}\end{aligned} \tag{21}$$

The time derivative of V is obtained as

$$\begin{aligned}\dot{V} &= 2\dot{V}_1+\mathbf{z}_1{}^{\mathrm{T}}\dot{\mathbf{z}}_1+2\dot{V}_3+\mathbf{z}_3{}^{\mathrm{T}}\dot{\mathbf{z}}_3\\ &= -2\mathbf{K}_1\mathbf{x}_{1e}^{\mathrm{T}}\mathbf{x}_{1e}-\mathbf{K}_2\mathbf{z}_1{}^{\mathrm{T}}\mathbf{z}_1-2\mathbf{K}_3\mathbf{x}_{3e}^{\mathrm{T}}\mathbf{x}_{3e}-\mathbf{K}_4\mathbf{z}_3{}^{\mathrm{T}}\mathbf{z}_3\end{aligned} \tag{22}$$

It is clear that the right side in (22) is negative definite because of the coefficient matrix $\mathbf{K}_i$ is positive, i.e., the entire system satisfies the Lyapunov stability theory.

3.4. Realization of control input

In section 3.1, we calculated the input command for translational motion. The absolute value of thrust and the target

value of the deflection angle are obtained from those commands. The thrust coordinate system is defined as shown in Fig.4.

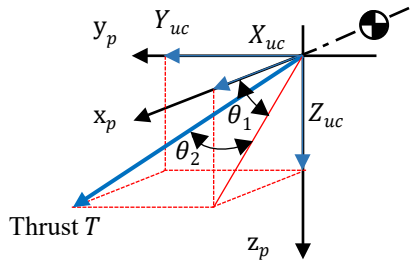


Figure 4: Definition of thrust coordinate system

The absolute value of thrust is obtained by (23).

$$T = \sqrt{X_{uc}^2 + Y_{uc}^2 + Z_{uc}^2} \tag{23}$$

X_{uc}, Y_{uc}, Z_{uc} are the commands of the input that is obtained to satisfy the ideal input of $\mathbf{F}_u$.The following inequalities denote the range of angles θ_1 and θ_2.

$$\begin{cases} -\pi \le \theta_1 \le \pi \\ -\pi/2 \le \theta_2 \le \pi/2 \end{cases}$$

where θ_1 and θ_2 are defined as

$$\tan\theta_1 = \frac{Z_{uc}}{X_{uc}}$$

$$\theta_1 = \tan^{-1}\left(\frac{Z_{uc}}{X_{uc}}\right) \tag{24}$$

$$\sin\theta_2 = \frac{Y_{uc}}{\sqrt{X_{uc}^2 + Z_{uc}^2}}$$

$$\theta_2 = \tan^{-1}\left(\frac{Y_{uc}}{\sqrt{X_{uc}^2 + Z_{uc}^2}}\right) \tag{25}$$

4. Numerical simulation

Numerical simulations are conducted to confirm the effect of the TVS on control performance in two cases: the level flight (Case 1) and the improvement of maneuverability by the TVS under the follow-up flight to the target altitude (Case 2) by numerical simulations. The conventional UAV without the TVS, its thrust command is obtained by using only X_{uc} from the translational controller. Moreover, regardless of the presence of the TVS, we introduce the first-order lag filter as written in (26) and (27) to the thrust, the TVS, and dynamic surface. T_m and T_s are the time constants.

$$\frac{1}{T_m s + 1} \tag{26}$$

$$\frac{1}{T_s s + 1} \tag{27}$$

Conditions of numerical simulation are shown in Tables 1 to 4.

Table 1: Specification of UAV

Full length L, m	0.823
Total height H, m	0.322
Wingspan b, m	1.40
MAC $\bar{c}_a$, m	0.200
Mass m, kg	0.220

Table 2: Initial condition and target value

Position $\mathbf{x_{1i}}$, m	$[0 \quad 0 \quad -10]^T$
Velocity $\mathbf{x_{2i}}$, m/s	$[5 \quad 0 \quad 0]^T$
Euler angle $\mathbf{x_{3i}}$, rad	$[0 \quad 0 \quad 0]^T$
Angular rate $\mathbf{x_{4i}}$, rad/s	$[0 \quad 0 \quad 0]^T$
Target position in Case 1 $\mathbf{x_{1c}}$, m	$[0 \quad 0 \quad -10]^T$
Target position in Case 2 $\mathbf{x_{1c}}$, m	$[0 \quad 0 \quad -12]^T$
Target velocity $\mathbf{x_{2c}}$, m/s	$[6 \quad 0 \quad 0]^T$

Table 3: Design control parameters

Translational controller gain $\mathbf{K_1}$	diag[1 1 1]
Translational controller gain $\mathbf{K_2}$	diag[0.5 0.5 0.5]
Rotational controller gain $\mathbf{K_3}$	diag[8 8 8]
Rotational controller gain $\mathbf{K_4}$	diag[10 10 10]

Table 4: Constraints on control input

Aileron angle δ_a, deg	± 45
Elevator angle δ_e, deg	± 45
Rudder angle δ_r, deg	± 30
Deflection angle θ_1, deg	± 135
Deflection angle θ_2, deg	± 90
Thrust T, N	0 or more
Time constant on thruster T_m, s	0.2
Time constant on dynamic surface and thrust vectoring system T_s, s	0.0076

In addition, the Dryden model [14] is used as a wind disturbance on numerical simulations. The wind disturbance profile is shown in Fig.5.

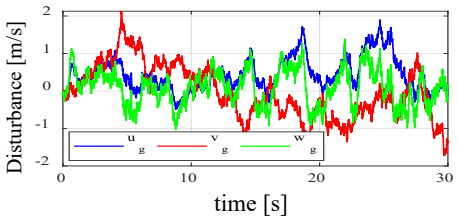
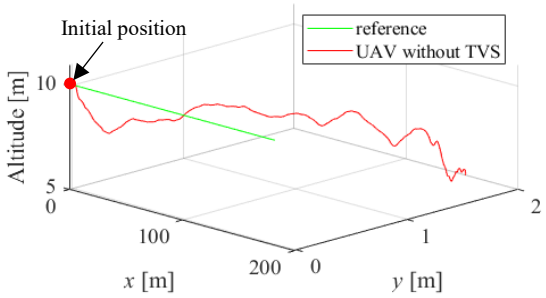


Figure 5: Wind disturbance

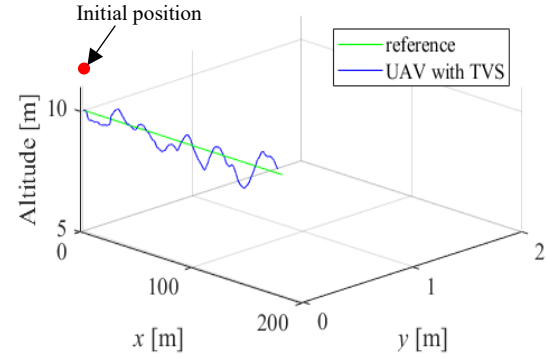
Fig.6 to Fig.9 show the numerical results in Case 1. It is clear that the UAV without the TVS cannot keep the altitude as shown in Figs.6(a), 6(b), and 6(d).

Furthermore, it should be noted that the TVS is also effective to control the UAV along y axis as shown in Fig.6(c). One of the main reasons for these flight is that the TVS can generate thrust in arbitrary direction as shown in Fig.8(a) under mechanical constraint. In fact, the inputs for translational motion in Fig.7(a) is constant because the UAV without the TVS cannot generate the thrust for the negative command as shown in Fig.7(b). Fig.7(c) and Fig.8(c) show the attitude angles of the UAV without and with

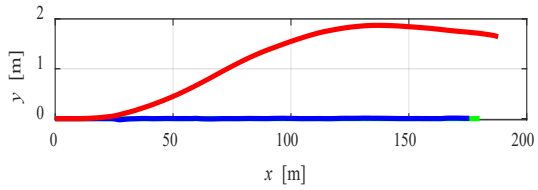
the TVS, respectively. It can be seen from these figures that the TVS suppressed the high-frequency vibration of pitch angle of the UAV. Fig.8(d) shows that the thrust change in Fig.8(a) leads to the sudden change of the elevator deflection to keep the altitude.



(a) In three-dimensional space (without TVS)

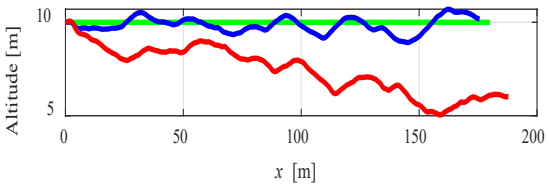


(b) In three-dimensional space (with TVS)



(c) In x-y plane

(–reference, –UAV with TVS, –UAV without TVS)



(d) Altitude corresponding to x

(–reference, –UAV withTVS, –UAV without TVS)

Figure 6: Trajectories of UAV in Case1

Fig.9 shows the phase portrait during the control of the UAV in Case 1. Using the proposed controller, the variables in the phase space follow those references, unlike the controller without the TVS.

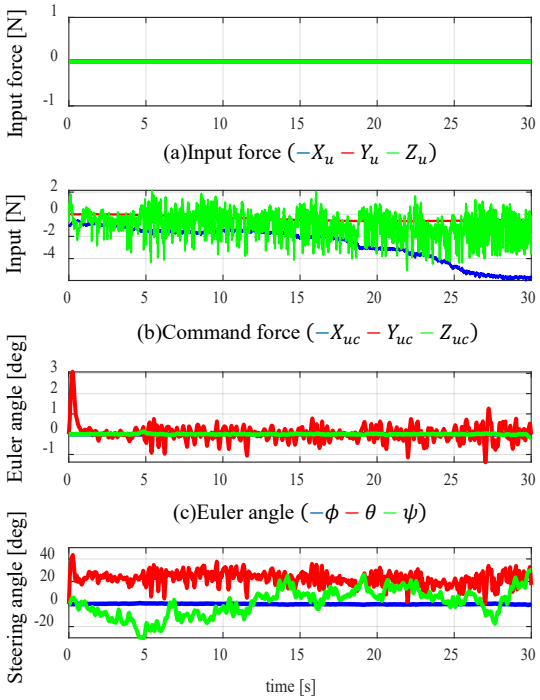


Figure7: Results in case of UAV without TVS (Case1)

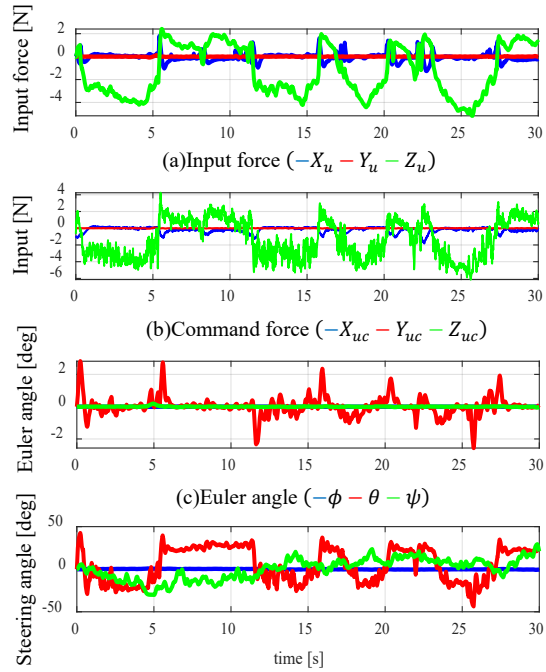


Figure 8: Results in case of UAV with TVS (Case1)

Fig.10 and Fig.11 show the numerical results in Case 2. The altitude of the UAV is controlled from the initial value 10 m to the final value 12 m as shown in Fig. 10(a). In this paper, the final value was decided by 20% higher than the initial value. However, there is no limitation to altitude in the proposed method, but it is necessary to be careful so that the altitude changes slightly. The attitude of the UAV with the TVS in Fig.10(c) is also controlled by control inputs in Fig.10(b) and Fig.10(d) similar to the Case 1.

Moreover, it can be seen in Fig.11 that the UAV follows the reference in the phase space.

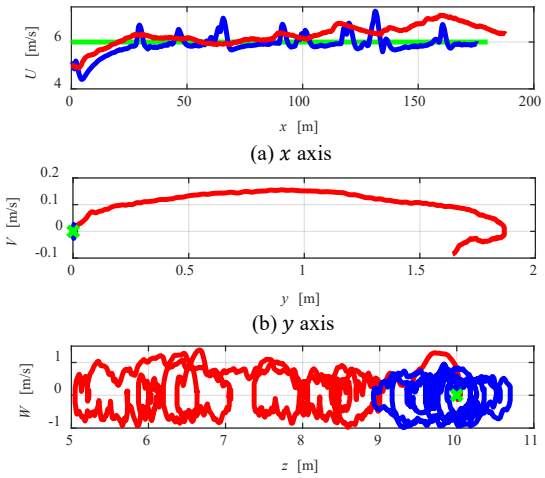
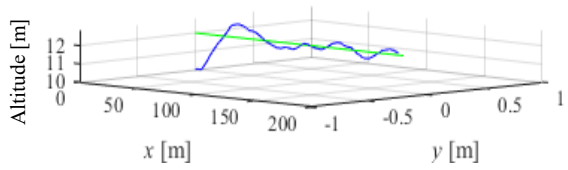
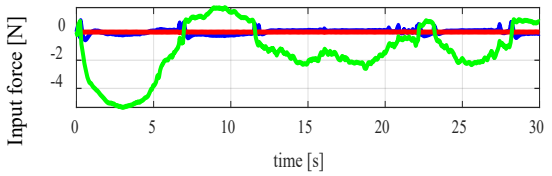


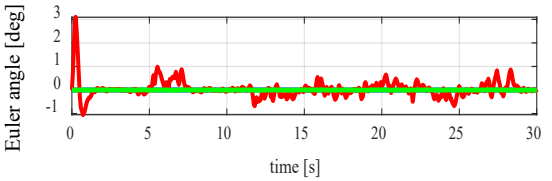
Figure 9: Phase portrait in Case1
(reference, UAV withTVS, UAV without TVS)



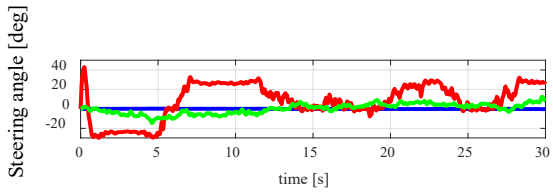
(a)Trajectory of UAV



(b)Input force $(-X_u - Y_u - Z_u)$



(c)Euler angle $(-\phi - \theta - \psi)$



(d)Input steering angle $(-\delta_a - \delta_e - \delta_r)$

Figure10: Results of the UAV with TVS in Case2

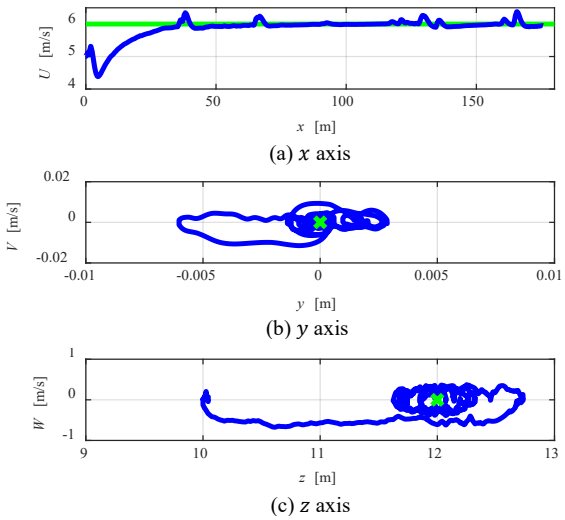


Figure11: Phase portrait in Case2
(reference, UAV withTVS.)

5. Conclusion

This paper proposed the nonlinear flight control system under disturbance for the UAV with a thrust vectoring system (TVS) using the backstepping method. The results of numerical simulation show that the TVS enables the stable flight even under disturbances for the two cases proposed.

The accurate estimation of the aerodynamic forces would afford the key to succeed the stable flight of the UAV with the TVS. In future work, we will verify the effectiveness of the proposed method experimentally.

Reference

[1] I V Kovalev, A A Voroshilova, and M V Karaseva, "Analysis of the current situation and development trend of the international cargo UAVs market," *J. Phys. Conf. Ser.*, 1399(5), 055095, 2019. https://doi.org/10.1088/1742-6596/1399/5/055095

[2] K. Shimizu, *Feedback Control Theory*. Corona Publishing Co., 2013.

[3] M. Krstic, I. Kanellakopoulos, and P. V. Kokotovic, *Nonlinear and Adaptive Control Design*. Wiley-Interscience, 1995.

[4] T. Espinoza, A. Dzul, R. Lozano, and P. Parada, "Backstepping — Sliding mode controllers applied to a fixed-wing UAV," in *2013 International Conference on Unmanned Aircraft Systems (ICUAS)*, Atlanta, GA, USA, 2013. https://doi.org/10.1109/ICUAS.2013.6564678

[5] A. Brezoescu, R. Lozano, and P. Castillo, "Bank to turn approach for airplane translational motion in unknown wind," in *2013 International Conference on Unmanned Aircraft Systems (ICUAS)*, Atlanta, GA,USA, 2013. https://doi.org/10.1109/ICUAS.2013.6564790

[6] K. Wu, B. Fan, and X. Zhang, "Trajectory following control of UAVs with wind disturbance," in *2017 36th Chinese Control Conference (CCC)*, Dalian, China, 2017. https://doi.org/10.23919/ChiCC.2017.8028144

[7] A. T. Espinoza Fraire, R. P. Parada Morado, A. E. Dzul López, and R. Lozano Leal, "Design and implementation of fixed-wing MAV controllers," in *2015 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED-UAS)*, Cancun, Mexico, 2015. https://doi.org/10.1109/RED-UAS.2015.7441004

[8] J. W. Roberts, R. Cory, and R. Tedrake, "On the controllability of fixed-wing perching," in *2009 American Control Conference*, St. Louis, MO, USA, 2009. https://doi.org/10.1109/ACC.2009.5160526

[9] K. Z. Y. Ang *et al.*, "Development of an unmanned tail-sitter with reconfigurable wings: U-Lion," in *11th IEEE International Conference on Control Automation (ICCA)*, Taichung, Taiwan, 2014. https://doi.org/10.1109/ICCA.2014.6871015

[10] C. Liang and C. Cai, "Modeling of a rotor/fixed-wing hybrid unmanned aerial vehicle," in *2017 36th Chinese Control Conference (CCC)*, Dalian, China, 2017. https://doi.org/10.23919/ChiCC.2017.8029181

[11] W. Wang, J. Zhu, M. Kuang, and X. Zhu, "Adaptive Attitude Control for a Tail-Sitter UAV with Single Thrust-Vectored Propeller," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, QLD, Australia, 2018. https://doi.org/10.1109/ICRA.2018.8463158

[12] M. Kuang, J. Zhu, W. Wang, and Y. Tang, "Flight controller design and demonstration of a thrust-vectored tailsitter," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, Singapore, 2017. https://doi.org/10.1109/ICRA.2017.7989605

[13] H. Kikkawa and K. Uchiyama, "Attitude control of a fixed-wing UAV using thrust vectoring system," in *2017 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED-UAS)*, Linkoping, Sweden, 2017. https://doi.org/10.1109/RED-UAS.2017.8101677

[14] Flying Qualities of Piloted Aircraft. Department of Defense Handbook. MIL-HDBK-1797. Washington, DC: U.S. Department of Defense, 1997.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 1, 285-293 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Implementation of Paraconsistent Logic Based PI Controller for TA Converter

Tapas Kumar Mohapatra*, Asim Kumar Dey, Krushna Keshab Mohapatra

SOA University, Department of Electrical Engineering, 751030, India

ARTICLE INFO

Article history:
Received: 14 October, 2019
Accepted: 23 January, 2020
Online: 10 February, 2020

Keywords:
Photo Voltaic (PV)
Paraconsistent Logic (PL)
Positive output voltage
TA converter
Paraconsistent logic based PI (PLPI)

ABSTRACT

A novel control technique is implemented for TA converter. This converter is used for charging of electric vehicle Battery from a Photo Voltaic (PV) Module. TA converter is one type of non-isolated buck-boost converter. This converter provides positive output voltage. This article is an extension work that is originally presented in 2019 IEEE International Conference on Sustainable Energy Technologies (ICSET). A Comparative study of the performances of different converters has been described. The Paraconsistent Logic (PL) is used for a PI controller. The performance of new Paraconsistent Logic based PI (PLPI) controller is compared with the conventional PI controller. This control has the similar action as PI and the logical actions as structured in Paraconsistent Logic. Operational details of TA converter with the different performances has been described in this article. MATLAB is used to validate the performance by developing both state space and simulation model. The robustness of the controller is compared with traditional PI controller. The hardware of the TA converter with the new PLPI control technique is projected and also matched with simulation results.

1 Introduction

Because of extraordinary utilization of petroleum derivative and air contamination, the auto mobile sectors are compelled to consider electric vehicles (EV). Presently a days EV battery modelling and its charging is an extraordinary research region. It is notable, switch mode power supply (SMPS) is the core of power transformation innovation. EV battery charging is promising through simple and efficient DC-DC converter. Mostly buck-boost converter is preferred for the EV battery charging process. Buck and boost converters are simple and efficient but unable to provide high voltage gain. Multiple circuits have been projected in different articles such as the complex topology developed by LUO [1]–[3] provides a high voltage gain. It is costly, big in size and losses is also more. Interleaved converters [4]–[10] are able to provide high voltage gain with less voltage stress but their operational mode ,converter arrangement and control strategies are very complex. Hwu and Peng [10] also developed buck-boost topology which can provide voltage gain of 2D with positive output. Quadratic converters [11] – [14] can provide high voltage gain with less efficiency. KY converter [15] has high voltage gain but increases complexity with additional switches and also increases the overall cost.Though the traditional buck-boost converter introduce high efficiency with reduced cost but unable to provide higher and positive output voltage. In 1991 Cuk suggested a converter, [16] which is able to provide a voltage gain of $\frac{D^2}{(1-D)^2}$ but can only operate in buck mode due to the presence of clamping diode D_1 and D_2. The converter which provides low gain can provide high or low voltage output at very high or very low duty ratio, which is practically difficult to achieve. A new converter proposed in [17], having little voltage ripple, negligible radio frequency intervention, and one shared ground switch but complexity increases due to the 7th circuit and also having different ground in its input and output terminals. In [18], a cascaded converter, combining two separated converters with both current source and sink, is functional for the thermoelectric generator. However, the voltage gain is also unnatural. Particularly, in order to attain high-voltage gain, the above converters must be functioning under extremely high or little duty cycle which is difficult to realize due to the practical limitations. Therefore, discovering new converter topology to overcome the negatives of the conventional ones for the requirements in industries is significant and appreciated.

TA converter [19] is proposed in the article which is one type non isolated positive output based buck-boost converter. It introduce a voltage gain of $\frac{D^2}{(1-D)^2}$. Though the voltage gain is similar to several other proposed converters but it disables few problems associated with others. This TA converter has several advantages over Shan and Faqiang converter [20] like:

*Tapas Kumar Mohapatra , Dept. of Electrical Engineering, SOA University & tapasmohapatra@soa.ac.in

- High voltage gain with positive output voltage.
- Less current stress in capacitor.
- Less voltage stress in inductor.
- Less source current for same load current in comparison to others.
- Can operate at CCM mode at low switching frequency.

The basic range of battery used for electric bike varies between 12 to 48 volt. In this article a battery of 12 volt, 7AH rating is used as a load, which is charged from a source varies between 6 to 30 volt. TA converter with PI controller is used to provide a constant voltage of 12 volt from a solar panel.

A controller/regulator is basically act as the brain of the system. Three types of conventional controller such as Proportional, Proportional and Integral (PI) and Proportional, Integral and Derivative (PID) have a wide scope of uses in SMPS application. With scientific advancement, controllers got computational assets about the 1980s, and turned out to be increasingly productive. The incorporation of micro controller chips for the implementation of controllers, in DC to DC converters, allowing the integration of algorithms with P, PI and PID activities, the expansion of structures with automatic tuning and massive simplicity in control parameter settings [21] - [24]. The primary goal of this work is to introduce the consequences of the PI controller using PL in the novel DC to DC boost converter and also compare the result with the conventional PI controller.

In section II of the article, the converter circuit diagram, operation and analysis is explained. Mathematical model and Simulink model is revealed in section III. In section IV, the Paraconsistent Logic based PI controller (PLPI) is explained. In section V hardware and simulation result of the converter is compared for PL-PI as well as the results compared with other converters. In the section VI the dynamic performance of the proposed controller is compared with conventional PI controller. Few comments and concluding remarks are given in section VII

2 Proposed TA converter Structure, Operation and analysis

2.1 Structure / Circuit Diagram

TA converter circuit diagram is shown in figure-1. The circuit elements of the converter are as follows: S_1 and S_2) - Power switches,(D_1 and D_2) – Diodes, (L_1 and L_1)- Inductors, (C_1 and C_1) - Capacitors. Time domain wave forms of the TA converter in CCM is shown in figure-3.

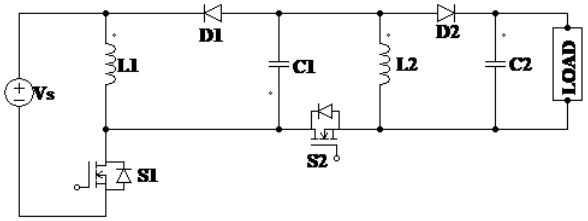


Figure 1: Circuit diagram of TA converter

2.2 Operating Principle and Analysis

In continuous conduction mode (CCM) the converter drives in two modes. First Mode -(0 < t < DT): In first mode S_1 and S_2 switches are simultaneously turned on and acts as short circuit, D_1 and D_2 diodes get reverse biased and act as open circuit. The direction of inductor currents are shown in figure - 2(a). C_2 provides the necessary output power to load. The following dynamic equations (1 - 8 will be obtained from figure - 2(a).

$$V_s - L_1\frac{di_{L1}}{dt} = 0 \quad (1)$$

$$\int_{I_{L1min}}^{I_{L1max}} di_{L1} = \int_0^{DT} \frac{V_s}{L_1} \quad (2)$$

$$\Delta I_{L1} = \frac{V_S}{L_1}DT \quad (3)$$

$$V_s - V_{L1} = 0 \quad (4)$$

$$V_{C1} - L_2\frac{di_{L2}}{dt} = 0 \quad (5)$$

$$\int_{I_{L2\min}}^{I_{L2\max}} di_{L2} = \int_0^{DT} \frac{V_{C1}}{L_2}dt \quad (6)$$

$$\Delta I_{L2} = \frac{V_{C1}}{L_2}DT \quad (7)$$

$$V_{C1} - V_{L1} = 0 \quad (8)$$

Second Mode - (0 < t < (1-D)T): In second mode S_1 and S_2 switches are simultaneously turned off and acts as open circuit, D_1 and D_2 diodes get forward biased by the inductors and act as short circuit. The direction of inductor currents are shown in Figure - 2(b). Inductor provides the necessary output power to load and to charge the capacitor C_2. The following dynamic equations (9 - 16) will be obtained from figure-2(b).

$$-V_{C1} - L_1\frac{di_{L1}}{dt} = 0 \quad (9)$$

$$\int_{I_{L1\max}}^{I_{L1\min}} di_{L1} = \int_0^{(1-D)T} \frac{-V_{C1}}{L_1}dt \quad (10)$$

$$\Delta I_{L1} = \frac{V_{C1}}{L_1}(1-D)T \quad (11)$$

$$-V_{C1} - V_{L1} = 0 \quad (12)$$

$$-V_{C2} - L_2\frac{di_{L2}}{dt} = 0 \quad (13)$$

$$\int_{I_{L2\max}}^{I_{L2\min}} di_{L2} = \int_0^{(1-D)T} \frac{-V_{C2}}{L_2}dt \quad (14)$$

$$\Delta I_{L2} = \frac{V_{C2}}{L_2}(1-D)T \quad (15)$$

$$-V_{C2} - V_{L2} = 0 \quad (16)$$

As we know the inductor voltage over a period is zero. We will obtain the following relations as given in equations (17-19). The inductor currents and capacitor voltages ripple are written in equation (20-23). Where f_{SW} is the switching frequency.

$$V_{C2} = V_0 = V_{C1}\frac{D}{1-D} \quad (17)$$

$$V_{C1} = V_S\frac{D}{1-D} \quad (18)$$

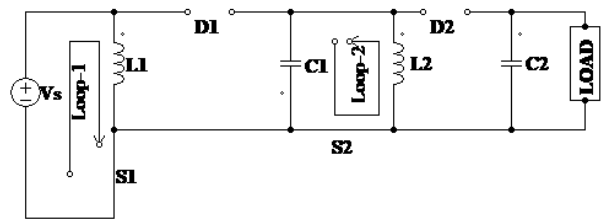
$$V_0 = V_S\frac{D^2}{(1-D)^2} \quad (19)$$

$$\Delta I_{L1} = \frac{V_S D}{L_1 f_{Sw}} \quad (20)$$

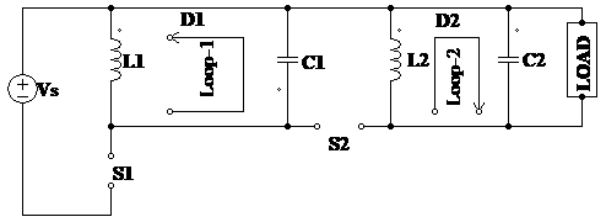
$$\Delta I_{L2} = \frac{V_S D^2}{L_2 f_{Sw}(1-D)} \quad (21)$$

$$\Delta V_{C1} = \frac{V_0 D}{RC_1 f_{Sw}(1-D)} \quad (22)$$

$$\Delta V_C 2 = \frac{V_0 D}{RC_2 f_{Sw}} \quad (23)$$



(a) First Mode



(b) Second Mode

Figure 2: Equivalent circuit diagram of two modes of TA converter

3 TA converter mathematical and simulink model

3.1 State Space Model

From the Mode-I and Mode-II the following dynamic equation will be obtained using which the state model can be obtained. The equations are given in (24-31).

$$\frac{di_{L1}}{dt} = \frac{V_S}{L_1} \quad (24) \qquad \frac{di_{L1}}{dt} = \frac{-V_{C1}}{L_1} \quad (28)$$

$$\frac{di_{L2}}{dt} = \frac{V_{C1}}{L_2} \quad (25) \qquad \frac{di_{L2}}{dt} = -\frac{V_{C2}}{L_2} \quad (29)$$

$$\frac{dV_{C2}}{dt} = -\frac{V_{C2}}{RC_2} \quad (26) \qquad \frac{dV_{C2}}{dt} = \frac{i_{L2}}{C_2} - \frac{V_{C2}}{RC_2} \quad (30)$$

$$\frac{dV_{C1}}{dt} = i_{L2} \quad (27) \qquad \frac{dV_{C1}}{dt} = -i_{L1} \quad (31)$$

$$\frac{di_{L1}}{dt} = D\frac{V_S}{L_1} - (1-D)\frac{V_{C1}}{L_1} \quad (32)$$

$$\frac{di_{L2}}{dt} = D\frac{V_{C1}}{L_2} - (1-D)\frac{V_{C2}}{L_2} \quad (33)$$

$$\frac{dV_{C2}}{dt} = -D\frac{V_{C2}}{RC_2} + (1-D)\frac{i_{L2}}{C_2} - (1-D)\frac{V_{C2}}{RC_2} \quad (34)$$

$$\frac{dV_{C1}}{dt} = Di_{L2} - (1-D)i_{L1} \quad (35)$$

The average model is given in equations (32-35).From the above equation it is clearly visible that the system is a fourth order one.

3.2 Simulation model and results

Figure-4 shows the Simulink model of the projected buck boost converter with PLPI controller. The circuit parameters that are used in the Simulink model are given in Table-II. Only one gate signal is essential for both the switches as both of them are operating simultaneously. A source of 18 volt is applied with output voltage of 40 volt in boost mode and 14 volt in buck mode. Input voltage is varied from 4volt to 18volt in boost mode and 18 volt to 100 volt in buck mode for the verification of robustness. Load resistance is also varied from 40Ω to 160Ω in boost and 5Ω to 50Ω in buck mode for the same verification. The load voltage remains as required with the above voltage and load resistance variations. In the figure -5 currents of each inductor and source, voltages of each capacitor, switch and diodes in boost mode are shown. Similarly in figure -6 the above waveforms in buck mode are displayed. It is clearly visible from figure - 5 and figure - 6 that i_{L1} never goes to negative value for both mode of operation.

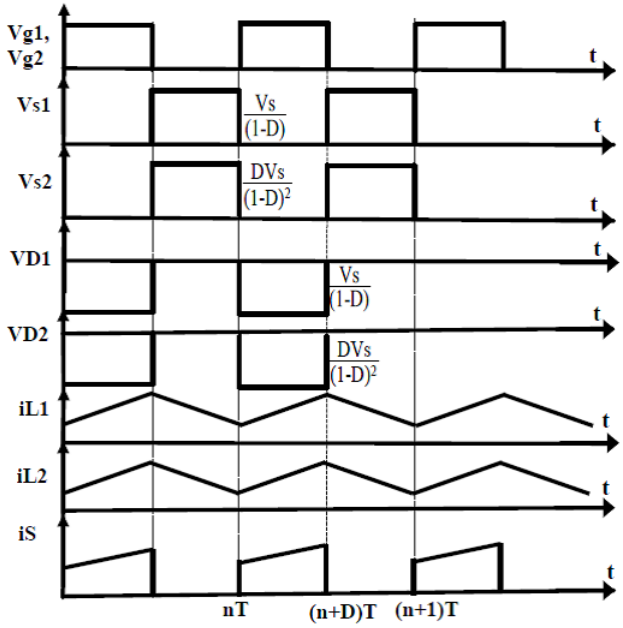


Figure 3: Different voltage and current wave forms

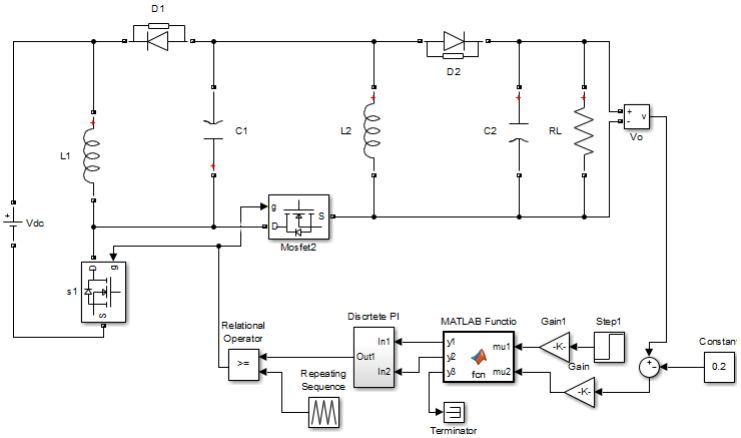


Figure 4: Matlab simulink model of proposed TA converter

4 Paraconsistent logic based PI controller

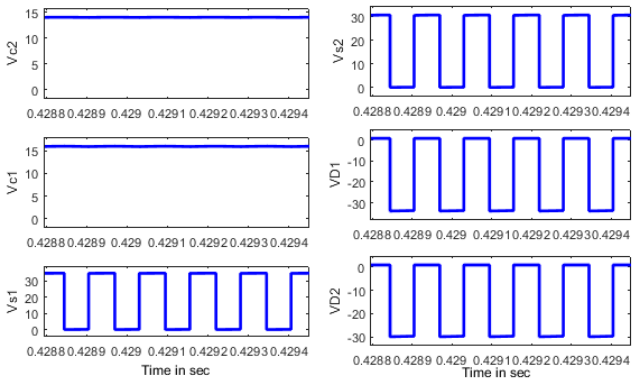
4.1 Logic of paraconsistent.

The conventional or Aristotelian logic underpins our present innovation is created on rigid binary laws and, along these lines, does not concede circumstances of redundancy, irregularities or those that are communicated by deficiency [25]. In conventional logic there is no logical inconsistency like something can't be both genuine and not valid in the meantime, when dealing with similar context. In any

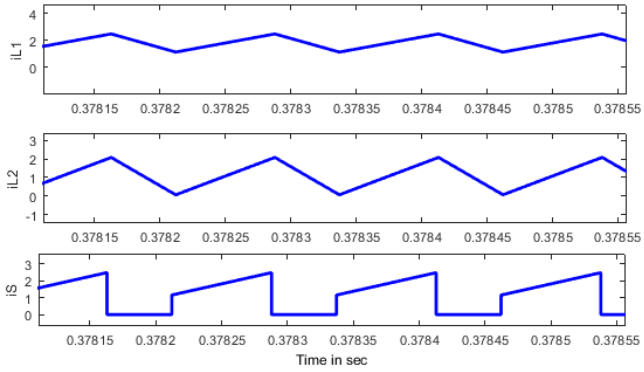
case, with the cut-off points of measures of control and eminence at present mandatory, progressively polished, make the philosophies of Conventional logic difficult to give effective control prototypes so improved for this mandate [25]. So as to act in circumstances where binary logic is difficult to be connected, different sorts of logic, called non-conventional, for example, Fuzzy, multivalued, paraconsistent, and so on, have as of late been made [25] - [26].

Non-conventional logics have the primary target is to restrict the binary principle of conventional logic and also catch diverse parts of casual conflicts. For instance, the principle of non-contradiction (PNC) in conventional logic explains that conflicting articulations can't both be valid in a similar sense in the meantime. Be that as it may, on account of signals as for a physical quantity beginning from two sensors, they might be conflicting, which gives conflicting data to make decision. In this way, control system will be progressively effective on the off chance that they are ready to act in circumstances where data might be conflicting.

A non-conventional Paraconsistent (PL) logic property states that it can tolerate the logical inconsistency in its establishments and is capable of managing conflicting signals. The PL has as its expansion the Paraconsistent Annotated Logic (PAL), which has a Hasse (Lattice FOUR) related with logical states existing at its vertices. Along these lines, sentences can be acquired where recommendations can be examined based on confirmations. In this PAL portrayal, the four extraordinary logic existing at the vertices of the PAL lattice are: Paracomplete ($\perp$), False (f), True (t) and Inconsistent ($\top$) [27] - [28].

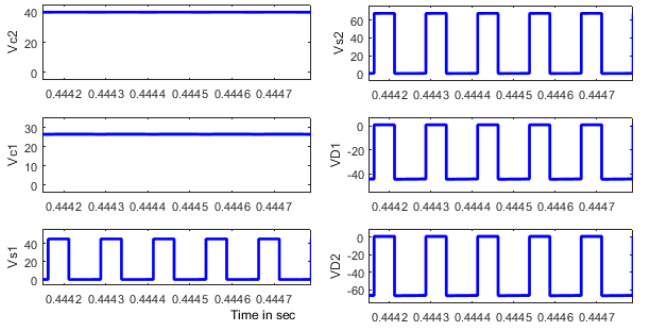


(a) $V_{c2}, V_{c1}, Vs1, V_{s2}, V_{D1}$ and V_{D2}

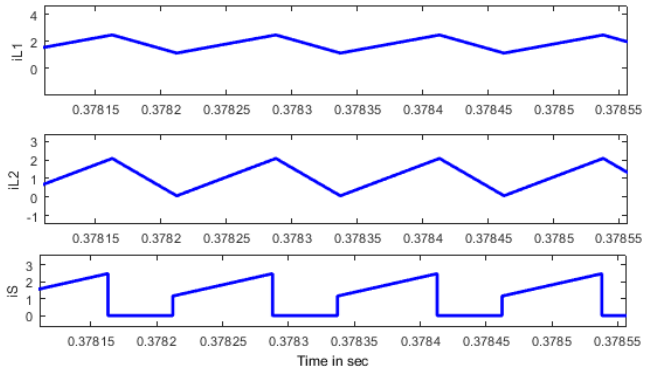


(b) $i_{L1}, i_{L2} and i_S$

Figure 6: Simulation result of TA converter in buck mode

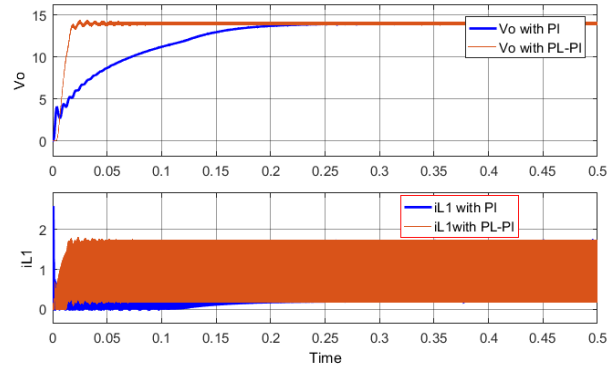


(a) $V_{c2}, V_{c1}, Vs1, V_{s2}, V_{D1}$ and V_{D2}

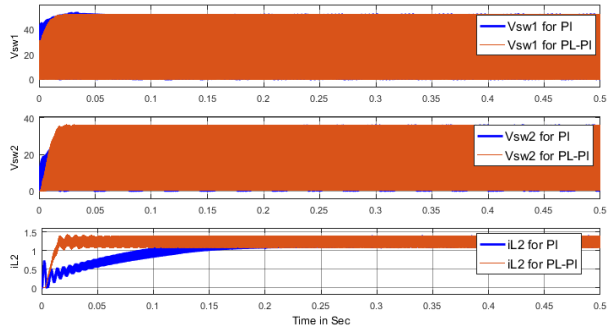


(b) $i_{L1}, i_{L2} and i_S$

Figure 5: Simulation result of TA converter in boost mode



(a) V_0 and i_{L1}



(b) $V_{sw1}, V_{sw2} and i_{L2}$

Figure 7: Comparison of results with PI and PL-PI

4.2 Logic with annotation of two values.

From a Lattice of four vertices it is conceivable to relate a sort of Paraconsistent Annotated Logic in which two evidential qualities related to a specific suggestion are considered. By translating the evidential values by methods for comment or standardized degrees of proof we can acquire conditions including the logical states that are characterized with degrees of proof from estimations. This sort of logic is called Paraconsistent Annotated Logic with explanation of two values (PAL2v) [27] - [28]. Figure-8(a) demonstrates the Lattice FOUR related at PAL2v.

In PAL2v an evidential nuclear equation of the shape P (μ , λ) can be viewed as an explanation for the recommendation p, where μ , $\lambda \in [0, 1]$ (genuine unit interim) [26] - [27]. Along these lines, the Evidence degree (μ) is an esteem that speaks to the proof ideal to the recommendation p, and the Evidence degree (λ) is an esteem that speaks to the proof adverse to the recommendation p. The relationship of a couple (μ ,λ) with a suggestion p implies that the level of ideal proof at p is μ , and the level of adverse proof at p is λ. As indicated by the comments in the related lattice appeared in figure-8(b)) [27]

- (1,1) → shows the presence of both, ideal and adverse proof totals, with a logical meaning of irregularity to the recommendation p.
- (1,0)→ shows the presence of complete ideal proof with adverse proof equal to zero, flagging an undertone of logical truth for the recommendation p.
- (0,0)→ shows the presence of ideal and adverse proof with the two values equal to zero, doling out a logical implication of uncertainty for the recommendation p.
- (0,1) → demonstrates the presence of ideal proof equal to zero and all out adverse proof, flagging a meaning of logical lie for the recommendation p.

The conditions of the PAL2v are acquired through a change where, at first, are viewed as characterised in a Unit Square in the Cartesian Plane (USCP) the level of ideal evidence μ in the y - axis, and the level of adverse evidence λ in the x - axis, as indicated by Figure-8(c)[26] - [28] A change that permits the degrees of proof of standardized qualities characterised in the x , y axis of the (USCP) to be situated on the X and Y axis of a four vertex PAL2v Lattice, given by equation - 36.

$$T(X, Y) = (xy, x + y - 1) \tag{36}$$

$$D_c = \mu - \lambda \tag{37}$$

$$D_{ct} = \mu + \lambda - 1 \tag{38}$$

$$\varepsilon\tau = (D_c D_{ct}) \tag{39}$$

$$\mu_{ctr} = (\mu + \lambda)/2 \tag{40}$$

$$d = \sqrt{(1 - |D_c|)^2 + (D_{ct})^2} \tag{41}$$

If $D_c > 0$

$$D_{CR} = 1 - \sqrt{(1 - |D_c|)^2 + (D_{ct})^2} \tag{42}$$

If $D_c < 0$

$$D_{CR} = \sqrt{(1 - |D_c|)^2 + (D_{ct})^2} - 1 \tag{43}$$

$$\mu_{ER} = (D_{CR} + 1)/2 \tag{44}$$

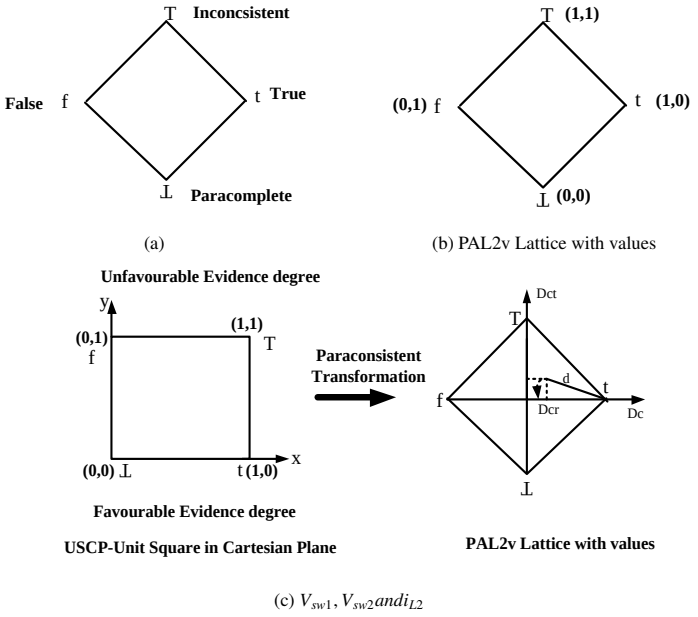


Figure 8: (a) Lattice associated with annotation of two values PAL2v. (b) PaL2v – Hasse finite Lattice with annotation. (c) Unit Square in the Cartesian Plane

Relating the parts of the change T(X , Y) concurring to the typical classification of PAL2v, where [26] - [28]: x = μ → Ideal Evidence degree, with $0 \leq \mu \leq 1$ and y = λ → level of Adverse Evidence, with $0 \leq \mu \leq 1$, where:

- The principal term acquired in the arranged match of the change condition is X = x-y = $\mu - \lambda$ → which will be known as the conviction degree D_c [27] - [29]. Accordingly, the conviction degree is acquired by equation - 37.
- The second term got in the arranged match of the change condition is: Y = x+y-1 = $\mu + \lambda$ - 1 → which will be known as the Contradiction Degree-Dct [27] - [29]. In this manner, the level of Contradiction is acquired by equation - 38.

The conditions of PAL2v permit the paraconsistent logical states to be found inside the lattice, the interpolation point is given by equation - 39. One can in this way settle on choices dependent on the vicinity of the logic state $\varepsilon\tau$ to the extraordinary logic states True (t) or False (f), found at the vertices of the PAL2v lattice. The Normalized Degree of Contradiction - μctr ,where a variety somewhere in the range of 0 and 1 is gotten, can be determined through condition [27] - [29] given in equation - 40. The Real Certainty degree D_{CR} is acquired by deciding the separation d in the PAL2v lattice as given in equation - 41. The (D_{CR}) values are determined by

the conditions appeared underneath [26] - [28] is given in equation - 42 and equation - 43 also linked by equation equation - 44.

The conditions gotten by the understandings made in the lattice related to PAL2v permit the making of algorithms utilized in examination and logical treatment of data signals. The Paraconsistent Analysis Node (PAN) algorithm can be utilized in a few fields of learning where fragmented and opposing data is dealt with fittingly through the PAL2v conditions. In this work, the PAN will be utilized to assemble a system of signals examination that embodies the control factors in a new converter.

4.3 Paraconsistent Algorithm

The explanation of a common PAN algorithm is specified under [26] - [28] State: P_1=input 1; P_2=input 2; O_1=outpu 1; O_2=output 2; O_3=output 3;

1 $\mu_1 = P_1$ (Ideal Evidence Degree), where $0 < \mu_1$. $\mu_2 = P_2$ (Ideal Evidence Degree), where $0 < \mu_2$.
$\lambda = 1 - \mu_2$ (Adverse Evidence Degree), where $0 < \lambda < 1$.

2 Evaluate the conviction degree $D_c = \mu 1 - \lambda$

3 Evaluate the degree of conviction $D_{ct} = \mu 1 + \lambda - 1$

4 Evaluate Normalized Degree of Contradiction. $\mu_{ctr} = (\mu 1 + \lambda)/2$

5 Evaluate the separation d in the PAL2v lattice. $d = \sqrt{(1 - |D_c|)^2 + (D_{ct})^2}$

6 Decide the output signals. If d > 1, then keep O_1=0.5; O_2= μ_{ctr} and $O_3 = D_{ct}$. Consider meaning less and go to exit. Else move to the next step

7 Decide actual conviction degree D_C. If $D_C > 0$ evaluate $D_{CR} = 1 - \sqrt{(1 - |D_c|)^2 + (D_{ct})^2}$
If $D_C < 0$ evaluate $D_{CR} = \sqrt{(1 - |D_c|)^2 + (D_{ct})^2} - 1$

8 Determine Resultant Real Evidence degree.
$\mu_{ER} = (D_{CR} + 1)/2$

9 The Real outputs are O_1=μ_{ER} , $O_2 = \mu_{ctr}$ and O_3= D_{ct} .

10 End.
In this article the Evidence degree was created as follows:
$0<\mu_2<1$ are proportional to $0.2V<P_2<14V$ and
$0<\mu_1<1$ are proportional to $0V<P_1<14V$.
$14V \rightarrow \mu_1 = 1$; with For $P_1 = 0V \rightarrow \mu_1 = 0$; and $P_1 = arelationof \mu_1 = P_1/14$.
$14V \rightarrow \mu_2 = 1$; with For $P_2 = 2V \rightarrow \mu_2 = 0$; and $P_2 = arelationof \mu_2 = (P_1 - 0.2)/14 and \lambda = 1 - \mu_2$.

5 Comparisons of Different Converters and the Control Technique

Table-I shows the comparisons of complexity, voltage gain, voltage and current stress, components of different converters such as traditional [30], Shan and Faqiang [20] and projected TA converter. Figure-9 curve shows the variation of voltage gain with the duty cycle. From observation it is clearly visible that TA converter provides a positive average inductor current but average inductor current under buck mode in Shan and Faqiang converter is negative. The peak inductor voltage V_{L2peak} and peak capacitor current I_{Cpeak} is less in TA converter. Figure-7 shows the comparison results of output voltage, inductor currents, switch voltages and source current between conventional PI and Paraconsistent logic based PI. From the figure it is clearly visible that the dynamic response of output voltage, inductor currents are faster in PL-PI than PI. It is also observed that for the same P and I gain the output voltage in PL-PI is obtained with less value of filter capacitor. The steady state values and time are same in both the controllers.

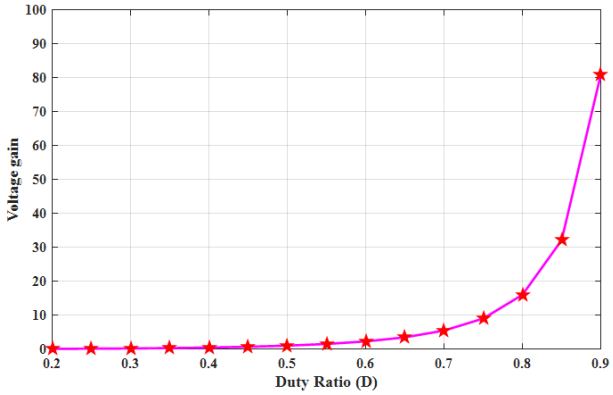


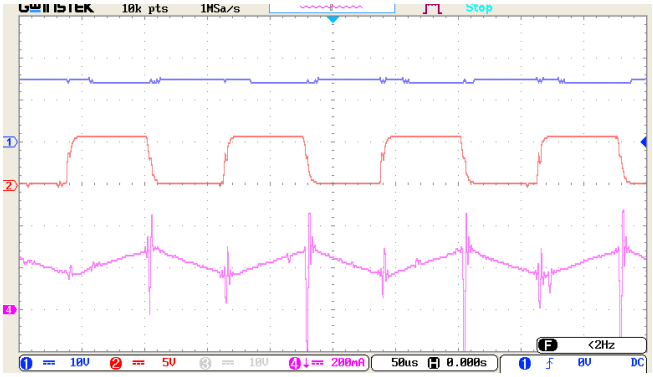
Figure 9: Voltage gain curve with duty ratio.

6 Hardware Parameters and Results

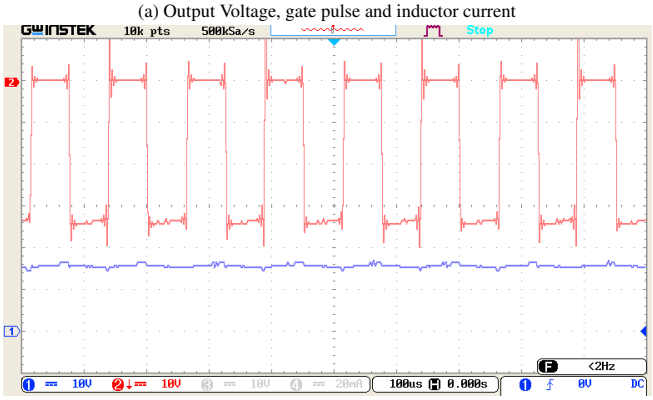
The parameters chosen for hardware implementation are provided in Table-II. PLPI controller is implemented for obtaining a fixed output voltage. The DSP (F28379D) Board is used for the designing of the proposed PLPI and traditional PI controller. The program is written in code composer studio and also the output file is burned on the DSP board using the same. Primarily a resistive load is taken to test the controller then the battery of rating 12volt 7Ah (Lead acid type) is taken as load. 1.4Amp of peak charging current is drawn by the battery. Figure-11 and Figure-10 shows the results in boost and buck mode respectively. Figure-12shows the photo of the complete hardware set up. Figure-13 which shows the efficiency curve explains clearly that efficiency in boost mode is higher than buck mode. The different results for battery load is shown in Figure-14.

7 Conclusion

A positive output based transformer less buck–boost (TA) converter is projected in this article. This converter overcomes several draw backs of the old-fashioned and Shan and Faqiang converter. The steady state operation principle, mathematical model, and performance comparisons with other converters are enlightened. The analysis is done using Matlab simulation and the verification of the same through experiment is done. This projected converter retains several merits in both boost and buck mode. This converter provides



(a) Output Voltage, gate pulse and inductor current

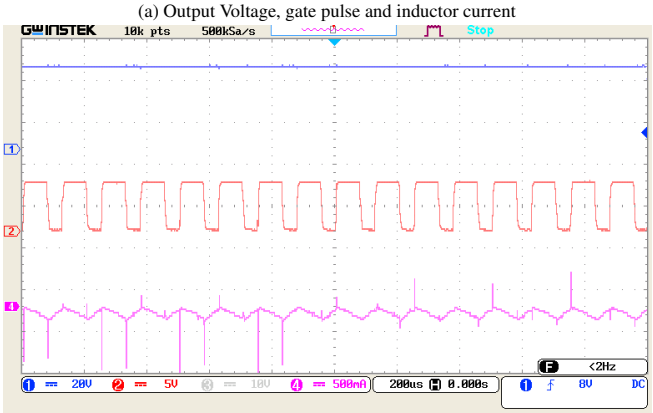


(b) Capacitor and Diode Voltage

Figure 10: Buck mode results.



(a) Output Voltage, gate pulse and inductor current



(b) Capacitor and Diode Voltage

Figure 11: Boost mode results.



Figure 12: Picture of the hardware setup of the proposed TA Converter with DSP and PC interface.

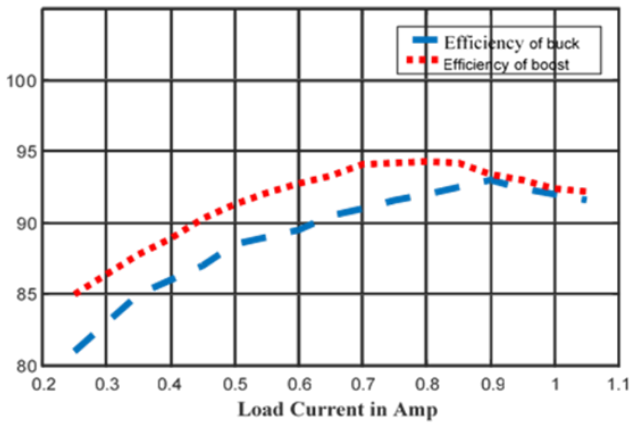


Figure 13: Efficiency Versus Load current for both Buck and Boost mode.



Figure 14: Output Voltage, gate pulse and inductor current for a Battery Load.

positive current through inductor and also provides positive output voltage. The converter control strategy is easy to implement. So, the planned converter is suitable for EV battery charging and industrial applications where high voltage gain is required. This article has also projected paraconsistent logic based PI controller. A comparison result is also shown in the figures. From the simulation and hardware result it is proved that the new control technique provides a better dynamic performance as well as provides better result with low value of capacitors.

References

[1] F. L. Luo and H. Ye, "Positive output cascade boost converters," Proc. Inst.Elect. Eng. Elect. Power Appl., vol. 151, no. 5, pp. 590–606, Sep 2004.

[2] Y. T. Chen,W. C. Lin, and R. H. Liang, "An interleaved high step-up DC–DC converter with double boost paths," Int. J. Circ. Theor. Appl., vol. 43, no. 8, pp. 967–983, Aug. 2015.

[3] L.W. Zhou, B. X. Zhu, Q. M. Luo, and S. Chen, "Interleaved non-isolated high step-up DC/DC converter based on the diode–capacitor multiplier,"IET Power Electron., vol. 7, no. 2, pp. 390–397, Feb. 2014.

[4] C. T. Pan, C. F. Chuang, and C. C. Chu, "A novel transformer less adaptable voltage quadrupler DC converter with low switch voltage stress," IEEE Trans. Power Electron., vol. 29, no. 9, pp. 4787–4796, Sep. 2014.

[5] K. I. Hwu, Member, IEEE, and Y. T. Yau, Student Member, IEEE" Two Types of KY Buck–Boost Converters" IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, VOL. 56, NO. 8, AUGUST 2009.

[6] D. Maksimovic and S. Cuk, "Switching converters with wide DC conversion range," IEEE Trans. Power Electron., vol. 6, no. 1, pp. 151–157, Jan.1991.

[7] Pan CT, Lai CM. A high-efficiency high step-up converter with low switch voltage stress for fuel-cell system applications. IEEE Transactions on Industrial Electronics. 2010 Jun;57(6):1998-2006.

[8] Luo FL. Six self-lift DC-DC converters, voltage lift technique. IEEE Transactions on Industrial Electronics. 2001 Dec;48(6):1268-72.

[9] Ye YM, Cheng KW. Quadratic boost converter with low buffer capacitor stress. IET Power Electronics. 2014 May;7(5):1162-70.

[10] K. I. Hwu and T. J. Peng, "A novel buck–boost converter combining KY and buck converters," IEEE Trans. Power Electron., vol. 27, no. 5, pp. 2236– 2241, May 2012.

[11] Morales-Saldana JA, Loera-Palomo R, Palacios-Hernández E, Gonzalez-Martinez JL. Modelling and control of a DC-DC quadratic boost converter with R 2 P 2. IET Power Electronics. 2014 Jan;7(1):11-22.

[12] Hua CC, Chiang HC, Chuang CW. New boost converter based on Sheppard-Taylor topology. IET Power Electronics. 2014 Jan;7(1):167-76.

[13] Kim RY, Lai JS. Aggregated modeling and control of a boost-buck cascade converter for maximum power point tracking of a thermoelectric generator. InApplied Power Electronics Conference and Exposition, 2008. APEC 2008. Twenty-Third Annual IEEE 2008 Feb 24 (pp. 1754-1760). IEEE.

[14] Li, Wuhua, Xiaodong Lv, Yan Deng, Jun Liu, and Xiangning He. "A review of non-isolated high step-up DC/DC converters in renewable energy applications." In Applied Power Electronics Conference and Exposition, 2009. APEC 2009. Twenty-Fourth Annual IEEE, pp. 364-369. IEEE, 2009.

[15] Yang LS, Liang TJ, Chen JF. Transformer less DC–DC converters with high step-up voltage gain. IEEE Transactions on Industrial Electronics. 2009 Aug;56(8):3144-52.

[16] Miao Z, Fang LL. Enhanced self-lift Cuk converter for negative-to-positive voltage conversion. IEEE transactions on power electronics. 2010 Sep;25(9):2227-33.

[17] Hua CC, Chiang HC, Chuang CW. New boost converter based on Sheppard-Taylor topology. IET Power Electronics. 2014 Jan;7(1):167-76.

[18] Ajami A, Ardi H, Farakhor A. Design, analysis and implementation of a buck–boost DC/DC converter. IET Power Electronics. 2014 Dec 1;7(12):2902-13.

[19] Mohapatra, Tapas Ku, Asim Ku Dey, and K. K. Mohapatra. "A Novel Nonisolated Positive Output TA Converter." 2019 IEEE International Conference on Sustainable Energy Technologies (ICSET). IEEE.

[20] Shan Miao, Faqiang Wang and Xikui Ma, "A New Transformer less Buck–Boost Converter with Positive Output Voltage", IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, VOL. 63, NO. 5, MAY 2016.

[21] Hwu KI, Yau YT. KY converter and its derivatives. IEEE Transactions on Power Electronics. 2009 Jan;24(1):128-37.

[22] Astrom KJ. T. haglund,"Auto Tuning of PID controllers," 1st ed. Research and triangle park, NC: instream. Soc. Amer. 1988.

[23] Zhao, D., Chai, T., and Wang H. Fu, J. (2014). Hybrid intelligent control for regrinding process in hematite beneficiation. Control Engineering Practice,22,217–230.

[24] Ogata, K. (2014). Modern control engineering (fifth ed.). Prentice Hall Inc, ISBN: 0-13-615673-8.

[25] Da Costa NC. On the theory of inconsistent formal systems. Notre dame journal of formal logic. 1974;15(4):497-510.

[26] Da Costa, N. C. A., and de Ronde, C. (2013). Found Phys, 43, 845. http://dx.doi.org/10.1007/ s10701-013-9721-9.

[27] Coelho MS, da Silva Filho JI, Côrtes HM, de Carvalho Jr A, Blos MF, Mario MC, Rocco A. Hybrid PI controller constructed with paraconsistent annotated logic. Control Engineering Practice. 2019 Mar 1;84:112-24.

[28] Baptista R, Da Silva Filho JI, Morilla JC. Proposal of Automatic Control three-way Valve through algorithms of Paraconsistent Annotated logic. Unisanta Science and Technology. 2013 Dec 14;2(2):69-73.

[29] Da Silva Filho JI. Algorithms based on paraconsistent annotated logic for applications in Expert systems. Expert System Software: Engineering, Advantages and Applications, Nova Science Publishers, Hauppauge. 2011:1-40.

[30] Y. He and F. L. Luo, "Analysis of Luo converters with voltage-lift circuit," Proc. Inst. Elect. Eng. Elect. Power Appl., vol. 152, no. 5, pp. 1239–1252, Sep. 2005.

Table 1: Comparison of different buck-boost converters

Parameters	Conventional Buck-Boost Converter [30]	Shan and Faqiang Converter [20]	TA converter [19]
Switches	1	2	2
Diodes	1	2	2
Capacitors	1	2	2
Inductors	1	2	2
Voltage gain	$\frac{D}{(1-D)}$	$(\frac{D}{(1-D)})^2$	$(\frac{D}{(1-D)})^2$
(Voltage stress on Switch-1)/V_S	$\frac{D}{(1-D)}$	$\frac{D}{(1-D)}$	$\frac{D}{(1-D)}$
(Voltage stress on Switch-2)/V_S		$(\frac{D}{(1-D)})^2$	$(\frac{D}{(1-D)})^2$
Nature of average Inductor current	Positive	Negative	Positive
complexity	2nd order	4th order	4th order
Peak Voltage of L_2	V_C	V_{C1}+V_{C2}	V_{C2} or V_{C1}
Peak current through C_1	I_{Lmax}	I_{L1max} + I_{L2max}	I_{L1max} or I_{2Lmax}

Table 2: Different parameters of TA [19] converter

Parameters	Values	Parameters	Values	Parameters	Values
Inductor L_1	1 mH	V_S	18 Volt	Diodes	1N
Inductor L_2	3.5 mH	Proportional gain	0.0145	Load Resistance for Buck	5-50ω
Capacitor C_1	1μF (PL-PI) and 100μF (PI)	Integral gain	39.346	Load resistance for Boost	40 - 160 ω
Capacitor C_2	1μF (PL-PI) and 220-1000μF (PI)	Switches	MOFET IRF	Battery	Lead Acid 12V, 7AH

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 5, 287-293 (2019)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

An Integrated Framework for Pronominal Anaphora Resolution in Malayalam

Ajees Arimbassery Pareed*, Sumam Mary Idicula

Department of Computer Science, CUSAT, 682022, India

ARTICLE INFO

Article history:
Received: 13 July, 2019
Accepted: 31 August, 2019
Online: 18 October, 2019

Keywords:
Malayalam
NLP
Anaphora resolution
Deep level tagger
Multilayer perceptron

ABSTRACT

Anaphora resolution is one of the old problems in Natural Language Processing. It is the process of identifying the antecedent of an anaphoric expression in a natural language text. Most of the NLP applications such as text summarization, question answering, information extraction, machine translation etc. require the successful resolution of anaphors. In this paper, we propose a methodology for the resolution of pronominal anaphors present in Malayalam text document. The proposed methodology is a hybrid architecture employing machine learning and rule-based techniques. In our study, we have used a deep level tagger developed using a machine learning based algorithm. The deep level tagger provides detailed information regarding the number and gender of nouns in a text document. The morphological features of the language are effectively utilized for the computational analysis of Malayalam text. Despite using less amount of linguistic features, our system provides better results which can be utilized for higher level NLP tasks such as question answering, text summarization, machine translation, etc.

1 Introduction

Anaphora is a Greek word that originated in 16th-century [1]. It is comprised of 'ana' and 'phora', where 'ana' means back, and 'phora' means carrying. Anaphors can successfully avoid the repetition of a word or phrase in a text document. They frequently occur in both written and spoken forms of the text. Resolving the correct antecedent of an anaphor from the set of possible antecedents is called anaphora resolution. It is one of the complicated problems in Natural Language Processing that has got considerable attention among NLP researchers. Halliday and Hassan defined anaphora as "the cohesion which points back to some previous item" [2]. According to Hirst, "anaphora is a device for making an abbreviated reference to some entity in the expectation that the receiver of the discourse will be able to dis-abbreviate the reference and, thereby, determine the identity of the entity" [3]. Hence, anaphora resolution is the process of disambiguating the antecedent of a referring expression from the set of entities in a discourse. A simple example illustrating the use of an anaphoric expression in English is given below.

Example: ***John*** went to school. ***He*** forgot to take his umbrella.

Here the word '***He***'is the anaphoric expression and '***John***'is the antecedent. Expressions referring to the actual entities in the preceding text are called anaphoric expressions. While the targets of anaphoric expressions are called antecedents.

It is a dream of humankind to communicate with the computers, since the beginning of the digital era. Communication with computers is only possible through a proper understanding of natural language text. The proper understanding of natural language text is possible only after anaphora resolution. There is a large amount of text data available online in Malayalam. Malayalam Wikipedia itself contains more than 30,000 articles. This warrants us to develop tools that can be used to explore digital information presented in Malayalam and other native languages. Anaphora resolution is one such tool that helps in semantic analysis of natural language text. It is also necessary for areas like text summarization, question answering, information retrieval, machine translation, etc.

Anaphora can be broadly classified into three categories [4]. They are pronominal anaphora, Definite noun phrase anaphora, and one anaphora. Among the various types of anaphora, pronominal anaphora is the most widely used one, which is realized through anaphoric pronouns. Pronouns are short words that can be substituted for a noun or noun phrase. They always refer to an entity that is already mentioned in the discourse. Pronominal anaphors are always stronger than full definite noun phrases [5]. They are again classified into five namely personal pronouns, possessive pronouns, reflexive pronouns, Demonstrative pronouns and relative pronouns. All the above-mentioned pronouns need not be anaphoric. A simple example illustrating the use of an anaphoric expression in Malayalam is given below.

*Department of Computer Science, CUSAT, 9061859697 & ajeesap87@gmail.com

രാജു വീട്ടിൽ പോയി. അവൻ പേനയെടുക്കാൻ മറന്നു.

Figure 1: Sample text illustrating the use of an anaphoric expression in Malayalam

Natural languages across the world are quite diverse in their structure. One language does not follow the syntax and semantics of the other. Moreover, anaphora resolution is a challenging task since it requires expertise in different domains of language processing such as lexical analysis, syntactic analysis, semantic analysis, discourse analysis, etc. When it comes to Malayalam, the major challenges associated with anaphora resolution are

- Lack of preprocessing tools
- Lack of standard datasets
- Agglutinative nature of the language
- Free word ordering behaviour
- Influence of cases in resolving the antecedent
- Use of different fonts to encode the text, etc

In this paper, we will focus only on resolution of pronominal anaphors present in a Malayalam text document using deep level tagging algorithms. The structure of this paper is as follows. Section 2 briefly reviews the related works. Section 3 explains the system architecture. Section 4 includes experiments and results. And section 5 concludes the article along with some directions for future work.

2 Related works

Most of the earlier works in anaphora resolution were reported in English and other Europian languages. 'STUDENT' is one of the earliest approaches to resolve anaphors through some automated mechanism. It was a high school algebra problem answering system attempted by Bobrow in 1964 [6]. He used a set of heuristics to resolve anaphors and other questions in a text. The second work in English was reported in 1972 by Winograd [7]. He was the first one to develop an algorithm to resolve pronominal anaphors. In his study, he considered previous noun phrases as the candidate antecedents for the pronominal anaphors. Syntactic positions examine all the possibilities from the preceding text. Wilks introduced a multilevel anaphora resolution system in 1973 [8]. Each level can resolve a set of anaphors based on the complexity to find the antecedent. The first level uses the knowledge of individual lexeme meaning to resolve anaphors. Similarly, the last level uses the "focus of attention" rules to find the topic of the sentence and hence the antecedent.

The modern era on anaphora resolution in English starts with the works of Hobbs in 1977 [9]. He developed an algorithm based on the syntactic constraints of the English language. Parse trees are used to search the antecedent of an anaphor in an optimal order such that the noun phrase upon which the search terminates is regarded as the probable antecedent of the pronoun. Carter reports a shallow processing approach in 1986 [10]. His work exploits the knowledge of syntax, semantics and local focusing. But he didn't give much attention to a large amount of domain knowledge, which he considered as a burden to process accurately. Grosz presented centring theory, a discourse-oriented study for anaphora resolution in 1977 [11]. Lappin and Leass reported another work for anaphora resolution in 1994 [12]. Their algorithm focuses on salience measures derived from the syntactic structure rather than the semantic factors present in the text. McCallum and Wellner presented a CRF based anaphora resolution system for English [13]. Similarly, Charniak and Elsner demonstrated an anaphora resolution system using Expectation Maximisation (EM) algorithm [14].

When it comes to Indian languages, works in anaphora resolution are reported only from few languages such as Hindi, Tamil, Bengali and Malayalam. This situation is contributed by the lack of standard parsers and sophisticated preprocessing tools for Indian languages. The first work in anaphora resolution for Malayalam was reported by Sopha et al. in 2000 as part of her doctoral thesis [15]. Their system (VASISTH) was the first multilingual system for anaphora resolution in Indian languages. Initially, the system was developed for Malayalam and then modified for Hindi. It explores the morphological richness of the language and makes limited use of syntax and other traditional linguistic features. Specific rules are designed for each type of anaphora separately. Athira et al. proposed an algorithm to resolve pronominal anaphors in Malayalam text [16]. They explored a hybrid system for the resolution of pronominal anaphors by incorporating rule-based and statistical approaches. The morphological richness of the language is also utilized in their study. Later, a pronominal anaphora resolution system was proposed by Sobha in 2007 [17]. In her work, she calculated salience weight for each candidate noun phrase from the set of possible candidates. Grammatical features determined the salience factors. The system could achieve comparable performance. High accuracy on a limited data set was the major drawback of her paper. Linguistic issues involved in Co-reference resolution in Malayalam was discussed in detail by Rajendran S in 2018.

3 Proposed method

Our goal is to build an anaphora resolution system for Malayalam which resolves all pronominal anaphors present in a text document. The architecture of the proposed system is shown in figure 2. It shows the general block diagram of pronominal anaphora resolution system for Malayalam. The first module of the architecture is the preprocessing module, where the sentence segmentation and word tokenization operations are carried out. After preprocessing, the preprocessed text is fed to the POS tagging module, where the words from the preprocessed input text are tagged with POS information. The tags used for this study belong to the BIS tagset. The third module of the architecture is the animate noun identification module which identifies the animate nouns from the set of nouns given by the POS tagging module. Animate nouns are the semantic category of nouns referring to persons and animals. They usually correlate with animate pronouns present in a natural language text. Hence identifying animate nouns from the set of nouns in a text document is necessary to resolve the pronominal anaphors present in the document.

After identifying the animate nouns in a POS-tagged text, they

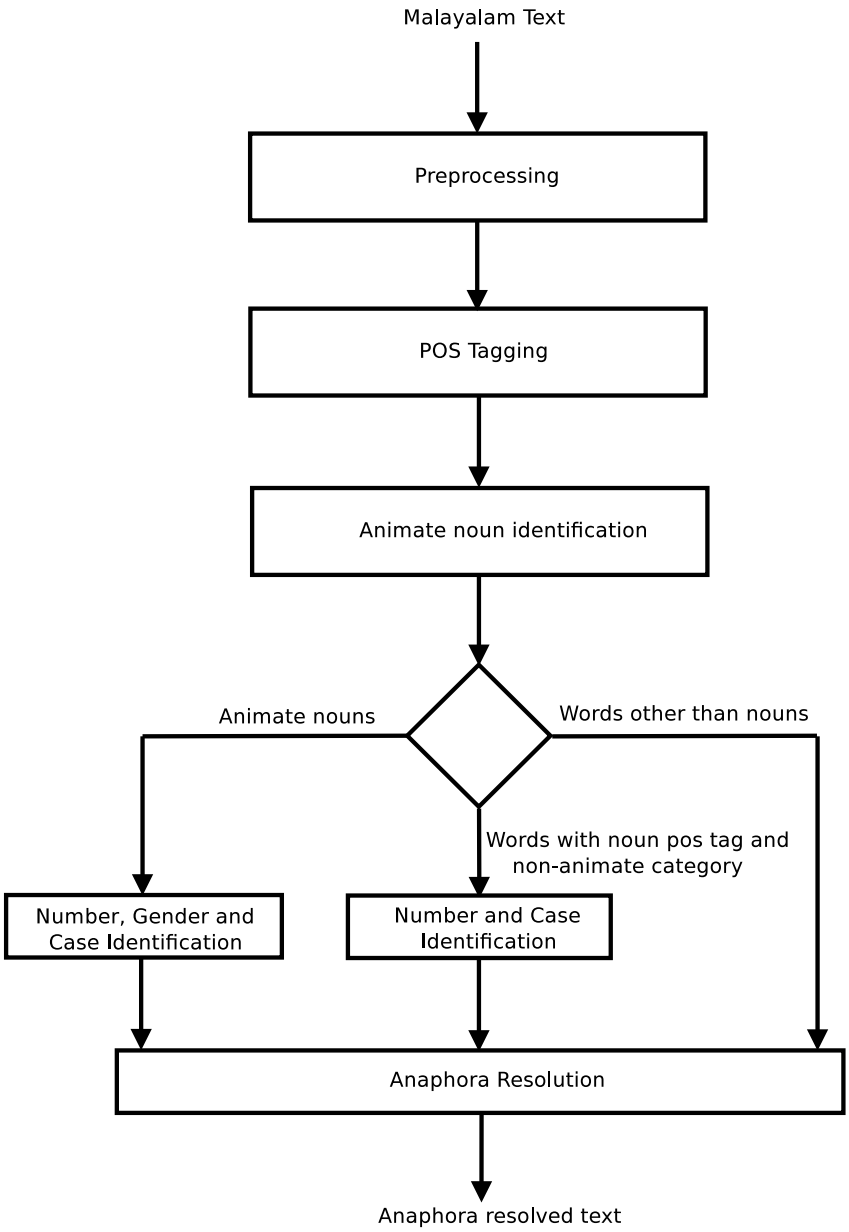


Figure 2: General architecture of the proposed system.

are sent to a number and gender identification neural network. Morphological features of the language are utilized to identify the number and gender information associated with animate nouns. The morphological features are captured with the help of a suffix stripper which can strip the suffixes of different length. Suffixes of different length act as the feature set for number and gender identification module. MLP classifier is used to build the in-depth analyzer of animate nouns. MLP is a class of feedforward artificial neural networks [18]. It is characterized by multiple hidden layers and non-linear activation. The case information associated with the animate nouns is also identified in this module. A rule-based approach is employed for this purpose.

The final module of the architecture is the anaphora resolution module, which is shown in figure 4. The deep level tagged text from the previous modules is the input towards the anaphora resolution module. For each pronoun present in the tagged text, a set of possible candidates are shortlisted. The possible candidates are animate nouns from the previous four sentences of the anaphor, that agree with the number and gender of the pronoun. After shortlisting the possible candidates, a set of salience values are assigned to each candidate. The salience factors are identified through proper analysis of Malayalam text. The set of salience factors we have considered in our study are given in table 2.

The clause information and subject-object information of shortlisted candidates are identified using a rule-based technique. For that, all the sentences containing the shortlisted candidates are divided into clauses. The words with a verb or adjective POS tag act as the clause boundary in a sentence [19]. In the case of adjectives, the nouns following them are also included in the same clause. For each clause identified in the last step, the subject, object and predicate

[['നരേന്ദ്രമോദി\N_NNP', 'സർക്കാരിനെ\N_NN', 'പിടിച്ചുലയ്ക്കുന്ന\V_VM_VNF', 'വിവാദബോംബാണ്\V_VAUX', 'വിജയ്\N_NNP', 'മല്യ\N_NNP', 'ഇന്നലെ\N_NN', 'ലണ്ടനിൽ\N_NNP', 'പൊട്ടിച്ചത്\V_VM_VNF', '.\RDPUNC'], ['ഇന്ത്യയിൽ\N_NNP', 'സാമ്പത്തിക\JJ', 'തട്ടിപ്പുകൾ\N_NN', 'നടത്തിയവർക്ക്\N_NN', 'ബിജെപിയുമായും\N_NN', 'കേന്ദ്ര\JJ', 'സർക്കാരിലെ\N_NN', 'ഉന്നതരുമായും\N_NN', 'അടുത്ത\N_NST', 'ബന്ധമുണ്ടെന്ന\V_VM_VNF', 'ആരോപണം\N_NN', 'പ്രതിപക്ഷ\N_NN', 'കക്ഷികൾ\N_NN', 'നിരന്തരം\RB', 'ഉയർത്തുന്ന\V_VM_VNF', 'ഘട്ടത്തിലാണ്\V_VAUX', 'അവർക്ക്\PR_PRP', 'വെടിക്കോപ്പ\N_NN', 'പകരുന്ന\V_VM_VNF', 'മല്യയുടെ\N_NNP', 'വെളിപ്പെടുത്തൽ-\N_NN', '.\RDPUNC'], ['ഇന്ത്യ\N_NNP', 'വിട്ടം\N_NN', 'മൻപ്\N_NN', 'ധനമന്ത്രി\N_NN', 'അരുൺ\N_NNP', 'ജയ്റ്റ്ലിയെ\N_NNP', 'കണ്ടിരുന്നു\V_VM_VF', '.\RDPUNC'], ['ചർച്ച\V_VM_VNF', 'നടത്തിയിരുന്നു\V_VM_VF', '.\RDPUNC']]



[['നരേന്ദ്രമോദി\N_NNP\ANI', 'സർക്കാരിനെ\N_NN\NANI', 'പിടിച്ചുലയ്ക്കുന്ന\V_VM_VNF', 'വിവാദബോംബാണ്\V_VAUX', 'വിജയ്\N_NNP\ANI', 'മല്യ\N_NNP\ANI', 'ഇന്നലെ\N_NN\NANI', 'ലണ്ടനിൽ\N_NNP\NANI', 'പൊട്ടിച്ചത്\V_VM_VNF', '.\RDPUNC'], ['ഇന്ത്യയിൽ\N_NNP\NANI', 'സാമ്പത്തിക\JJ', 'തട്ടിപ്പുകൾ\N_NN\NANI', 'നടത്തിയവർക്ക്\N_NN\NANI', 'ബിജെപിയുമായും\N_NN\NANI', 'കേന്ദ്ര\JJ', 'സർക്കാരിലെ\N_NN\NANI', 'ഉന്നതരുമായും\N_NN\NANI', 'അടുത്ത\N_NST', 'ബന്ധമുണ്ടെന്ന\V_VM_VNF', 'ആരോപണം\N_NN\NANI', 'പ്രതിപക്ഷ\N_NN\NANI', 'കക്ഷികൾ\N_NN\NANI', 'നിരന്തരം\RB', 'ഉയർത്തുന്ന\V_VM_VNF', 'ഘട്ടത്തിലാണ്\V_VAUX', 'അവർക്ക്\PR_PRP', 'വെടിക്കോപ്പ\N_NN\NANI', 'പകരുന്ന\V_VM_VNF', 'മല്യയുടെ\N_NNP\ANI', 'വെളിപ്പെടുത്തൽ-\N_NN\NANI', '.\RDPUNC'], ['ഇന്ത്യ\N_NNP\NANI', 'വിട്ടം\N_NN\NANI', 'മൻപ്\N_NN\NANI', 'ധനമന്ത്രി\N_NN\NANI', 'അരുൺ\N_NNP\ANI', 'ജയ്റ്റ്ലിയെ\N_NNP\ANI', 'കണ്ടിരുന്നു\V_VM_VF', '.\RDPUNC'], ['ചർച്ച\V_VM_VNF', 'നടത്തിയിരുന്നു\V_VM_VF', '.\RDPUNC']]

Figure 3: Sample text showing the input and output of the Animate Noun Identification module

Table 1: Distinct classes considered for deep level tagging

Label	Class
M/S	Singular male
M/P	Plural male
F/S	Singular female
F/P	Plural female

information are extracted using the following set of rules [20].

- Predicate —The verb or adjective present in the clause will form the predicate.
- Object —If the predicate is a verb, the noun preceding it will be the object and if the predicate is an adjective, the noun following it will be the object.
- Subject —If the predicate is a verb, the noun preceding the object will be the subject. And if there is no noun preceding the object in the same clause, follow the given rules.
    1. If the predicate of the preceding clause is an adjective, its object will form the subject.
    2. Else the subject of the preceding clause will be the subject.

  Conversely, if the predicate is an adjective, the noun preceding it will be the subject.

Finally, the possible candidates are sorted based on the total salience score and the candidate having the highest salience score is selected as the antecedent of the anaphor.

4 Experiments and Results

In this section, we will discuss the experiments on each phase of the architecture one by one. The first phase is the preprocessing of raw Malayalam text. We have used NLTK implementation sentence tokenizer and word tokenizer for this purpose [21]. Sentence tokenizer converts raw Malayalam text into a list of sentences, and word tokenizer converts each sentence into a list of words. The preprocessed text is provided to the POS tagging module, where the tagged text is generated. A CRF based tagger developed in our department is employed for this purpose [22]. The third module of the architecture deals with animate noun identification. MLP classifier is used to build the animate noun identifier. A set of nouns belonging to both animate and non-animate category is used as the training data. Each noun in the training data is labelled with corresponding class information (animate/non-animate). A set of 109430 nouns were prepared in this way and used for building the machine learning model. 'Fasttext'is used to convert words into vectors. 'Fasttext'is an efficient way to convert words to vectors developed by Facebook's AI research lab [23].

The different configuration of the neural network with various word vector sizes was attempted to build the animate noun identification model. A four-layered architecture with two hidden layers seemed to be the optimum one. The first layer is the input layer which accepts the input. Second and third layers are hidden lay-

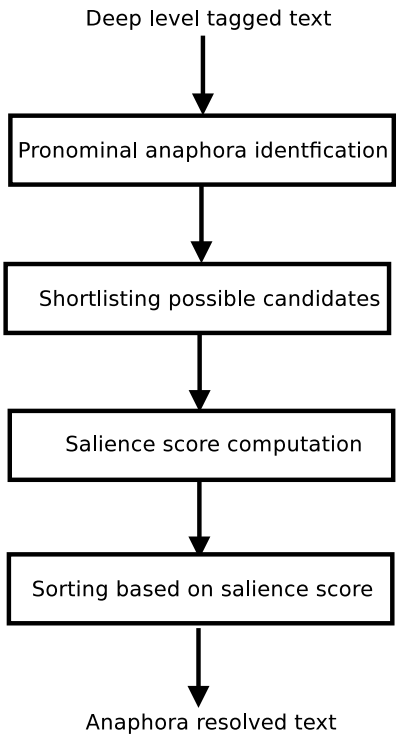


Figure 4: Architecture of anaphora resolution module

Table 2: Salince factors and their weights

Salience factors	Weights
Current sentence	100
Preceding sentences	Reduce sentence score by 10
Subject	80
Object	50
Case	10 to 40
Current clause	100
Immediate clause	50

ers with 'tanh' activation. The last layer is the output layer which generates the output. Softmax is used as the activation function in the output layer. A network with more hidden layers could not improve the performance and only delayed the convergence time. Hence we finalized our hidden layer size as (2,200), where 2 is the number of hidden layers, and 200 is the size of the hidden layers. The maximum accuracy obtained by the classifier on test data is 90.01%. Figure 3 shows the input and output of the animate noun identification module on sample text.

The fourth module of the architecture deals with deep level tagging of animate nouns. MLP classifier is used to build the deep level tagging module. Training data required for the study is a set of names with different number and gender features. Distinct classes considered for the investigation are shown in table 1. 12600 nouns belonging to different categories were prepared for this purpose. Suffixes of various length are used as the feature set for the classifier. Since words and suffixes are symbolic units, it can't be directly fed into neural networks. Hence, we converted our feature set numeric values using Dictvectorizer, a python functionality [24]. Various configurations of the MLP classifier are attempted in our study. A smaller network was not able to represent the data efficiently and increasing the network size did not improve accuracy. Hence, we experimentally finalized our hidden layer size as (2,200), where 2 is the number of hidden layers, and 200 is the size of each hidden layer. 'Relu' is used as the activation function and 'Adam' as the optimizer. The maximum accuracy was obtained by the classifier when the number of features was eight. The performance of the number and gender classifier with a different number of features is shown in figure 5.

In order to evaluate the performance of the proposed algorithm, experiments were conducted on two different types of datasets. These experiments were aimed to find the contribution of animate noun identifier and deep level tagger of nouns to the overall accuracy of the system.

4.1 Dataset1

This dataset contains short stories from children story domain. We have taken short stories in the Malayalam language from [25], a popular site for Malayalam short stories. Each story contains around 20 to 30 sentences of moderate length. Since children stories follow a

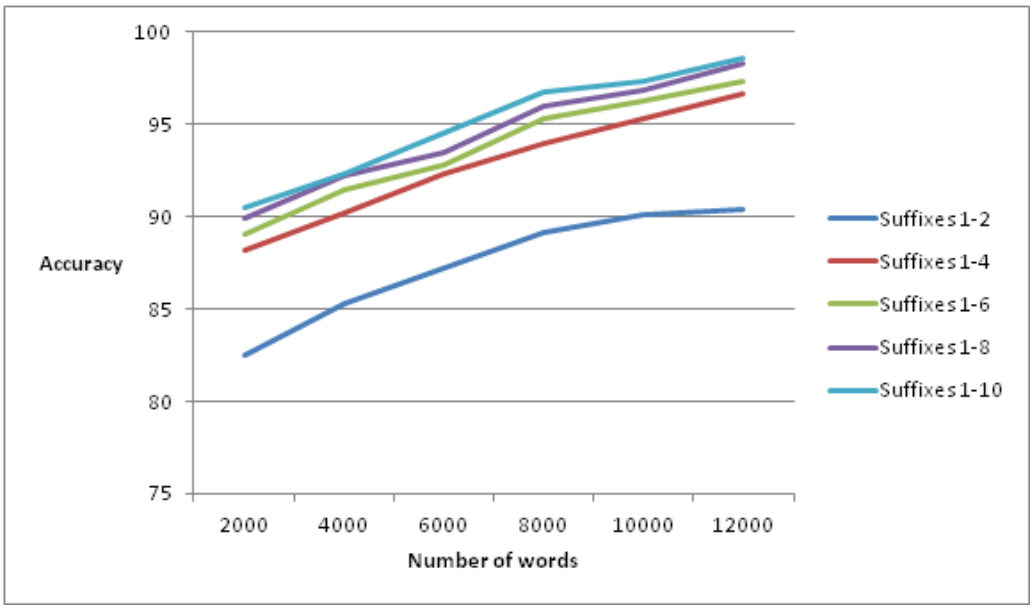


Figure 5: Effect of different features on the performance of the number and gender classifier

Table 3: Results of our experiments on dataset 1

Data set	Total sentences	Total pronouns	Correctly resolved	Accuracy
1	353	19	16	84%
2	466	21	17	80%

straight forward narrative style, the structural complexity of the sentences will be limited. The performance of the anaphora resolution algorithm on this dataset can represent the baseline performance.

4.2 Dataset2

This dataset contains text from news article domain. We have taken news articles from [26], a popular site for Malayalam news. Unlike the children short stories, here the structural complexity of the sentences is not limited. The length of the sentences may go over 25 words. Each article contains around 30 to 40 sentences. Results of our algorithm on both the datasets are summarized in table 3.

From the above experiments, it is observed that the proposed system achieves an overall accuracy of 82.5%. The correctness of the results obtained is verified with the help of language experts. Malayalam is a free word order language, which affects the accuracy of the system. It is observed that the case information associated with the nouns can also decide the antecedent of the pronouns. Some pronouns refer to both animate and inanimate things, which also affects the performance of the system.

5 Conclusion

This paper presents the experimental results of anaphora resolution system in Malayalam language using a hybrid approach. The exclusive feature of the proposed system is the use of machine learning as well as the rule-based technique for the resolution of pronominal anaphors. Malayalam is an agglutinative free word order language, and hence it has many complications in resolving pronouns. The deeper level analysis of nouns helps in the semantic understanding of natural language text. The power of word embedding is also exploited in the study, which helped in recognizing the animate nouns from non-animate nouns. Experimental results show that the deep level analysis of nouns is essential for resolving the pronominal anaphors present in a text document. The morphological richness of Malayalam language is also utilized in this study with the help of suffix stripping algorithms. In our experiments, we have observed that the increase in morphological features results in the increase in accuracy of tagging systems. Hence, incorporating morphological features in the analysis of natural language text appears to be promising for languages such as Malayalam. In the current work, we have focussed only on the resolution of pronominal anaphora. In future, we aim at the resolution of non-pronominal anaphora such as event anaphora, one-anaphora, etc.

References

[1] Kath. Word of the day – anaphora. 2018 (accessed December 7, 2018).

[2] MAK Halliday and R Hasan. 1976: Cohesion in english. london: Longman. 1976.

[3] Grahame Hirst. Anaphora in natural language understanding. 1981.

[4] M Sadanandam and D Chandra Mohan. Telugu pronominal anaphora resolution.

[5] Dan Jurafsky and James H Martin. *Speech and language processing*, volume 3. Pearson London, 2014.

[6] Daniel G Bobrow. A question-answering system for high school algebra word problems. In *Proceedings of the October 27-29, 1964, fall joint computer conference, part I*, pages 591–614. ACM, 1964.

[7] Terry Winograd. Understanding natural language. *Cognitive psychology*, 3(1):1–191, 1972.

[8] Yorick Wilks. Preference semantics. Technical report, STANFORD UNIV CA DEPT OF COMPUTER SCIENCE, 1973.

[9] Jerry R Hobbs. Pronoun resolution. *ACM SIGART Bulletin*, (61):28–28, 1977.

[10] David Maclean Carter. *A shallow processing approach to anaphor resolution*. PhD thesis, University of Cambridge, 1986.

[11] Barbara J Grosz. The representation and use of focus in dialogue understanding. Technical report, SRI INTERNATIONAL MENLO PARK CA MENLO PARK United States, 1977.

[12] Shalom Lappin and Herbert J Leass. An algorithm for pronominal anaphora resolution. *Computational linguistics*, 20(4):535–561, 1994.

[13] Ben Wellner, Andrew McCallum, Fuchun Peng, and Michael Hay. An integrated, conditional model of information extraction and coreference with application to citation matching. In *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, pages 593–601. AUAI Press, 2004.

[14] Eugene Charniak and Micha Elsner. Em works for pronoun anaphora resolution. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pages 148–156. Association for Computational Linguistics, 2009.

[15] L Sobha and BN Patnaik. Vasisth: An anaphora resolution system for indian languages. In *Proceedings of International Conference on Artificial and Computational Intelligence for Decision, Control and Automation in Engineering and Industrial Applications*, 2000.

[16] S Athira, TS Lekshmi, RR Rajeev, Elizabeth Sherly, and PC Reghuraj. Pronominal anaphora resolution using salience score for malayalam. In *2014 First International Conference on Computational Systems and Communications (ICCSC)*, pages 47–51. IEEE, 2014.

[17] L Sobha. Resolution of pronominals in tamil. In *2007 International Conference on Computing: Theory and Applications (ICCTA'07)*, pages 475–479. IEEE, 2007.

[18] Pedregosa Fabian, Varoquaux Gaël, Gramfort Alexandre, Michel Vincent, Thirion Bertrand, Grisel Olivier, Blondel Mathieu, Prettenhofer Peter, Dubourg Vincent, Vanderplas Jake, et al. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[19] MS Bindu and Sumam Mary Idicula. A hybrid model for phrase chunking employing artificial immunity system and rule based methods. *International Journal of Artificial Intelligence & Applications*, 2(4):95, 2011.

[20] Rajina Kabeer and Sumam Mary Idicula. Text summarization for malayalam documents—an experience. In *2014 International Conference on Data Science & Engineering (ICDSE)*, pages 145–150. IEEE, 2014.

[21] Edward Loper and Steven Bird. Nltk: The natural language toolkit. In *Proceedings of the ACL-02 Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics-Volume 1*, pages 63–70. Association for Computational Linguistics, 2002.

[22] AP Ajees and Sumam Mary Idicula. A pos tagger for malayalam using conditional random fields.

[23] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146, 2017.

[24] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct):2825–2830, 2011.

[25] Bed time stories - wonderful stories in malayalam for kids. `https://kuttykadhakal.blogspot.com/`. Accessed: 2018-09-30.

[26] Madhyamam. `https://www.madhyamam.com/`. Accessed: 2018-08-28.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 291-296 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Fuzzy Modelling using Firefly Algorithm for Phishing Detection

Noor Syahirah Nordin[1], Mohd Arfian Ismail[1,*], Vitaliy Mezhuyev[2], Shahreen Kasim[3], Mohd Saberi Mohamad[4,5], Ashraf Osman Ibrahim[6]

[1] *Faculty of Computing, College of Computing and Applied Sciences, Universiti Malaysia Pahang, 26300, Malaysia.*

[2] *FH JOANNEUM University of Applied Sciences, Institute of Industrial Management, Austria.*

[3] *Soft Computing and Data Mining Centre, Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn, Johor, Malaysia.*

[4] *Institute for Artificial Intelligence and Big Data, Universiti Malaysia Kelantan, City Campus, Pengkalan Chepa, 16100 Kota Bharu, Kelantan, Malaysia.*

[5] *Faculty of Bioengineering and Technology, Universiti Malaysia Kelantan, Jeli Campus, Lock Bag 100, 17600 Jeli, Kelantan, Malaysia.*

[6] *Faculty of computer Science and Information Technology, Alzaiem Alazhari University, Khartoum North 13311, Sudan.*

ARTICLE INFO

Article history:
Received: 30 August, 2019
Accepted: 16 November, 2019
Online: 12 December, 2019

Keywords:
Fuzzy modeling
Firefly algorithm
Phishing detection

ABSTRACT

A fuzzy system is a rule-based system that uses human experts' knowledge to make a particular decision, while fuzzy modeling refers to the identification process of the fuzzy parameters. To generate the fuzzy parameters automatically, an optimization method is needed. One of the suitable methods provides the Firefly Algorithm (FA). FA is a nature-inspired algorithm that uses fireflies' behavior to interpret data. This study explains in detail how fuzzy modeling works by using FA for detecting phishing. Phishing is an unsettled security problem that occurs in the world of internet connected computers. In order to experiment with the proposed method for the security threats, a database of phishing websites and SMS from different sources were used. As a result, the average accuracy for the phishing websites dataset achieved 98.86%, while the average value for the SMS dataset is 97.49%. In conclusion, both datasets show the best result in terms of the accuracy value for fuzzy modeling by using FA.

1. Introduction

Phishing is a cyber-attack criminal activity that is intended to steal sensitive information such as credit card information or account login credential from users by using bogus websites [1], [2]. There are three components in phishing techniques; medium of phishing, vector to transmit the attack, and technical approaches used during the attack. The first component, the medium of phishing is the base means of conveying the phishing attacks to the victims which involve three bases; internet, voice, and short messaging service (SMS). The second component, the vector that defines the vehicle in place for launching the attack such as Email, eFax, websites, and social networks that are accessible through the Internet. The last component is the technical approaches which are used to improve the phishing effectiveness during an attack.

Nowadays, many approaches are being used by the phishers to steal personal information such as browser vulnerabilities, mobile phone or man-in-them-middle. A phishing attack is the simplest kind of security threat, but at the same time is the most effective and dangerous violence. This is due to the fact that attackers use malware to remotely control a victim's device for their particular intention such as spying or stealing personal information. While not many people are aware that they may be a victim of a phishing attack, it is poses a major threat in network security. Therefore, many researchers are focused on the methods to detect phishing efficiently and produce better results than the previous methods.

*Corresponding Author: Mohd Arfian Ismail, arfian@ump.edu.my

One way to improve the efficiency to detect the phishing attack is by using the fuzzy techniques. With the fuzzy system, people can make an intelligent decision that works based on the combination of several factors. However, this method is a time consuming and does not guarantee an optimum solution because a fuzzy system requires the identification of the fuzzy parameters; fuzzy rules and the membership function. Hence, the optimization method needs to be applied in the system in order to tune the parameters of the fuzzy system automatically. This paper proposes a novel method by application of the Firefly Algorithm (FA). FA can be considered as a recent optimization method that is being used in artificial intelligence [3].

2. Materials and Methods

2.1. Related Works

There are many existing phishing detection techniques proposed in recent years. Researchers [4] proposed a software named anti-phishing simulator that collects phishing and spam messages where the users can examine the link addresses in the mail. It prescribes whether the messages can be classified as a phishing attack by using the Bayesian classification algorithm. Authors [5] have given a phishing site detection approach via URL analyses. Their work uses a URL detection method to discover phishing websites using a random forest algorithm. They have limited the feature set of URL detection to eight out of thirty-one features. The parameters considered to measure the accuracy level include f-measure, ROC Curve, precision, and sensitivity for analysis purposes. As a result, the accuracy level of this method was 95% [5]. By using the machine learning, researchers [6] proposes a hybrid solution that combines three approaches; blacklist and whitelist, heuristics, and visual similarity. The hybrid solution will be fed to the machine learning algorithm to calculate the accuracy results. By using different approaches , it will produce better accuracy and provide more efficient protection system [6]. According to [7], PhishBox is a new approach for phishing validation and detection that collects phishing data in real-time. The modules in this method include extract-transform-load, modelling, voting, monitoring, and visualization. The results [7] show that the proposed method has achieved high performance compared to the other works. Authors [8] used the C4.5 decision tree algorithm to analyse phishing sites. The data contain URL heuristics and the sites ranked to decide on a phishing attack. The proposed method extracts URL features and calculates their heuristic value. Then, the C4.5 decision tree algorithm was used to generate rules and identify the probability of phishing. There are 9 features of the URL used in the proposed method to detect phishing sites. The results showed that the method is more robust and precise compared to previous methods. Researchers [9] proposed a secured methodology for anti-phishing. The algorithm and techniques used were balanced block replacement, advanced encryption standard, and code generation. There are two phases in the proposed techniques which are user registration phase and user login phase. The method [9] is not flexible to accommodate the increasing number of consumers, therefore it will be difficult to provide a unique code to each user. Authors [10] proposed phishing emails detection using Cuckoo Search SVM (CS-SVM). Cuckoo Search algorithm was used for parameter selection. It extracts 23 features that are used to construct the hybrid classifier. This method uses the measures of a true positive rate, a false positive rate, and an accuracy as evaluation metric to evaluate the performance. CS-SVM shows a 91 percent higher result in terms of phishing email detection accuracy at different training sets when compared with traditional SVM classifier [10].

Overall, a lot of methods have been proposed by other researchers and have their advantages and disadvantages in producing the results. For detecting phishing in real-time, the researchers create new software that uses specific tools and algorithms to collect phishing data. On the other hand, the machine learning approach was recognised by researchers as the most effective method.

2.2. Fuzzy System

The fuzzy system is a rule-based system which works by using the fuzzy logic to reason data. Most of the fuzzy concepts come from the human language. It is an approach based on the "degrees of truth" and it imitates the way humans make decisions that involve the possibilities between YES and NO.

To ensure that the fuzzy system works properly, fuzzy parameters are needed. The parameters are the fuzzy rules and membership functions [11]. Fuzzy rules were originally obtained from human experts through the knowledge engineering processes. However, this approach cannot be applied when there are no human experts or if the data are too complex. Besides, membership function is a function that defines the degree to which a given input belongs, where the output is between 0 and 1. To implement a fuzzy logic technique , four elements are required which are fuzzification, fuzzy inference engine, fuzzy rule base, and defuzzification. The elements of the fuzzy system are shown in Figure 1 while the list and description of the components in a fuzzy system are described in Table 1.

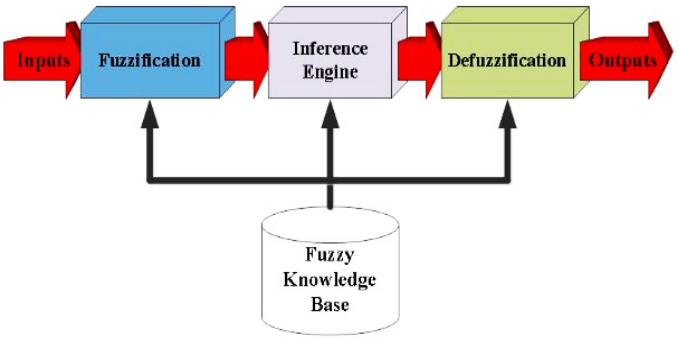


Figure 1: Elements of the fuzzy system

Table 1: Description of the fuzzy system elements

Component	Description
Fuzzy knowledge base	It contains a set of fuzzy sets and fuzzy rules
Fuzzification	Convert crisp data into the membership function
Inference Engine	Perform fuzzy operation by combining membership functions with the fuzzy rules to obtain the fuzzy output
Defuzzification	Convert fuzzy output into crisp data

Meanwhile, fuzzy modelling is a task of finding or identifying the fuzzy parameters to achieve the desired behaviour. An effective method should be used to generate the fuzzy parameters automatically from data. Regarding that, using an optimization method is the best choice by automatically generating the fuzzy parameters from available data [12].

2.3. Representation of Fuzzy Parameter

In the study, the fuzzy rule is defined by using numerical form where it is responsible for representing the fuzzy sets in the model. The fuzzy rules use the IF-THEN form as shown in Figure 2.

IF (A_j^i IS R_j^i) AND (A_j^i IS R_j^i) AND ….. (A_j^i IS R_j^i) THEN (CL IS C^i)

Where,

A = linguistic variable (attribute of the data)
R = linguistic value (value of the data)
CL = class attribute of the data
C = output of the linguistic value
j = number of attribute in the data
i = number of rules that is used

Figure 2: IF-THEN rule form description

The A in the IF-THEN form represents the input of the linguistic value while the output is the value of C for the class variable. The value of the attribute (A) is set in a range of 0 until 3, where the values of 1 and 2 are represented as phishing and legitimate respectively while the values of 0 and 3 do not apply. It gives meaning where if the fuzzy rule produces the value of 0 or 3 in the attribute's value, the attribute will not be included as the fuzzy set. Meanwhile, the value of class (CL) is set to 0 and 1, where it represents the result of the attribute.

The example of the process of encoding the fuzzy rules in the case study shown in Figure 3. The sample was taken from the Phishing Websites Dataset. In this example, almost all 30 attributes were included in the fuzzy rule in which 8 produced the value of 0, which is not included in the fuzzy sets. This is because the attribute has a value of 1; "*phishing*". If the fuzzy set stated "*legitimate*", the value will be 2. At the end of the fuzzy rules that represent a membership function, the results will be shown.

The schematic shape of the membership function used in this study is trapezoidal. The encoding process of the membership function is demonstrated in Figure 4, where it applied on the first dataset (the phishing websites dataset). The value of every parameter represented the starting point of the overlap in the membership functions. Since the number of attributes in the dataset is 31, the length of the membership function is equal to 31 as well.

2.4. Firefly Algorithm

The FA is an algorithm that was developed by Xin-She Yang at Cambridge University in 2007 [13]. The FA is a swarm intelligence-based metaheuristic approach inspired by the behaviour of fireflies. The flashing light of fireflies acts as a signal system or communication to attract other fireflies. It can also function as a protective warning mechanism. The flashing characteristics of the fireflies are as follows: i) All fireflies are unisex, therefore they become attracted to other fireflies without being concerned about their sex; ii) The less bright fireflies will move to the other fireflies who has the brighter flash as the attractiveness is proportional to their brightness. The attractiveness and brightness of the firefly are reduced as the distance increase. If no firefly is brighter than them, the fireflies will move randomly without the right direction; iii) The brightness of a firefly is determined by the setting of the objective function to be optimized.

Rule 1: IF (*have IP* IS *phishing*) AND (*URL length* IS *phishing*) AND (*shortening service* IS *phishing*) AND (*at symbol* IS *phishing*) AND (*double slash* IS *phishing*) AND (*prefix suffix* IS *phishing*) AND (*sub domain* IS *phishing*) AND (*SSL final* IS *phishing*) AND (*domain registration length* IS *phishing*) AND (*favicon* IS *phishing*) AND (*port* IS *phishing*) AND (*https* IS *phishing*) AND (*request URL* IS *phishing*) AND (*anchor* IS *phishing*) AND (*link in tags* IS *phishing*) AND (*SFH* IS *phishing*) AND (*redirect* IS *phishing*) AND (*on mouse over* IS *phishing*) AND (*right click* IS *phishing*) AND (*pop up window* IS *phishing*) AND (*iframe* IS *phishing*) then (*website* IS *phishing*)

A_1^1	A_2^1	A_3^1	A_4^1	A_5^1	A_6^1	A_7^1	A_8^1	A_9^1	A_{10}^1	A_{11}^1	A_{12}^1	A_{13}^1	A_{14}^1	A_{15}^1	A_{16}^1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

A_{17}^1	A_{18}^1	A_{19}^1	A_{20}^1	A_{21}^1	A_{22}^1	A_{23}^1	A_{24}^1	A_{25}^1	A_{26}^1	A_{27}^1	A_{28}^1	A_{29}^1	A_{30}^1	CL
0	0	1	1	1	1	1	0	0	0	0	0	0	0	1

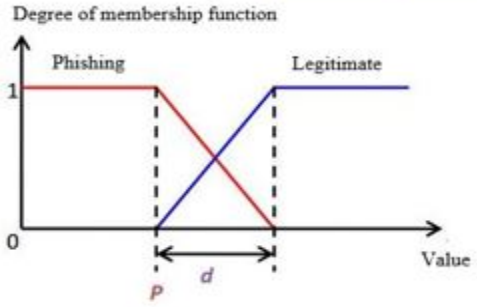
Where,

A_1^1	:	Have IP	A_{11}^1	:	Port	A_{21}^1	:	Right click
A_2^1	:	URL length	A_{12}^1	:	Https	A_{21}^1	:	Pop up window
A_3^1	:	Shortening service	A_{13}^1	:	Request URL	A_{23}^1	:	Iframe
A_4^1	:	At symbol	A_{14}^1	:	URL of anchor	A_{24}^1	:	Age of domain
A_5^1	:	Double slash	A_{15}^1	:	Links in tags	A_{25}^1	:	DNS record
A_6^1	:	Prefix	A_{16}^1	:	SFH	A_{26}^1	:	Web traffic
A_7^1	:	Sub domain	A_{17}^1	:	Submit to email	A_{27}^1	:	Page rank
A_8^1	:	SSL	A_{18}^1	:	Abnormal URL	A_{28}^1	:	Google index
A_9^1	:	Domain registration length	A_{19}^1	:	Redirect	A_{29}^1	:	Links pointing to page
A_{10}^1	:	Favicon	A_{20}^1	:	On mouse over	A_{30}^1	:	Statistical report

CL : Phishing / Legitimate

Parameter	Attribute
A_1^1	Have IP
A_2^1	URL length
A_3^1	Shortening service
A_4^1	At symbol
A_5^1	Double slash
A_6^1	Prefix
A_7^1	Sub domain
A_8^1	SSL
A_9^1	Domain registration length
A_{10}^1	Favicon
A_{11}^1	Port
A_{12}^1	Https
A_{13}^1	Request URL
A_{14}^1	URL of anchor
A_{15}^1	Links in tags
A_{16}^1	SFH

Parameter	Attribute
A_{17}^1	Submit to email
A_{18}^1	Abnormal URL
A_{19}^1	Redirect
A_{20}^1	On mouse over
A_{21}^1	Right click
A_{22}^1	Pop up window
A_{23}^1	Iframe
A_{24}^1	Age of domain
A_{25}^1	DNS record
A_{26}^1	Web traffic
A_{27}^1	Page rank
A_{28}^1	Google index
A_{29}^1	Links pointing to page
A_{30}^1	Statistical report
CL	Phishing / Legitimate



	A_1^1	A_2^1	A_3^1	A_4^1	A_5^1	A_6^1	A_7^1	A_8^1	A_9^1	A_{10}^1	A_{11}^1	A_{12}^1	A_{13}^1	A_{14}^1	A_{15}^1	A_{16}^1
p	1	0	1	0	0	2	1	0	0	0	2	2	1	1	0	1
d	0	0	1	0	1	1	1	0	0	1	2	1	1	1	0	2

	A_{17}^1	A_{18}^1	A_{19}^1	A_{20}^1	A_{21}^1	A_{22}^1	A_{23}^1	A_{24}^1	A_{25}^1	A_{26}^1	A_{27}^1	A_{28}^1	A_{29}^1	A_{30}^1	CL
p	1	1	0	2	2	1	0	1	0	0	2	0	1	1	0
d	0	0	1	1	1	1	1	2	2	1	2	2	1	0	0

The work of FA is started by initializing the objective function, followed by generating the initial population of fireflies. Then continued with determining the light intensity and ranking the fireflies before updating the fireflies' position in the population. Figure 5 shows the flowchart of the FA.

2.5. Model and Experimental Data

In order to test the efficiency of the proposed method in detecting phishing, two benchmarks datasets were used. The first dataset is taken from the University of California, Irvine (UCI) machine learning repository. This database is the trusted and most widely used dataset for detecting phishing attacks, which can be accessed at http://archive.ics.uci.edu/ml/. The dataset involved is a phishing website dataset that contains 2456 instances and 30 attributes. The dataset collected is mainly from trusted sources; PhishTank archive, MillerSmiles archive, Google searching operators.

Meanwhile, the second dataset that was used contains SMS messages that were obtained from the Unicamp website at http://www.dt.fee.unicamp.br/~tiago/smsspamcollection/. The dataset contains 5574 instances and 2 attributes that were collected from various sources such as Grumbletext Web, NUS SMS Corpus, and SMS Spam Corpus v.0.1 Big.

The experiments were conducted by using 10-fold cross-validation for both datasets. The cross-validation was performed by partitioning the data into 10 partitions where every partition consists of equal number of data in it. The process was then repeated 10 times. In evaluating these experiments, the results were measured by their fitness value which is equivalent to the accuracy of the model. There are three categories considered in the model, which are the best solution, worst solution, and average value. The best solution is the highest accuracy in each experiment amongst the 10-fold cross-validation in a single run while the worst solution indicates the lowest accuracy value in a single run. The purpose of determining the best and worst result in every experiment is to show the potential of fuzzy modelling by using FA. Meanwhile, the average solution is the mean value of all the fitness value results in the experiments.

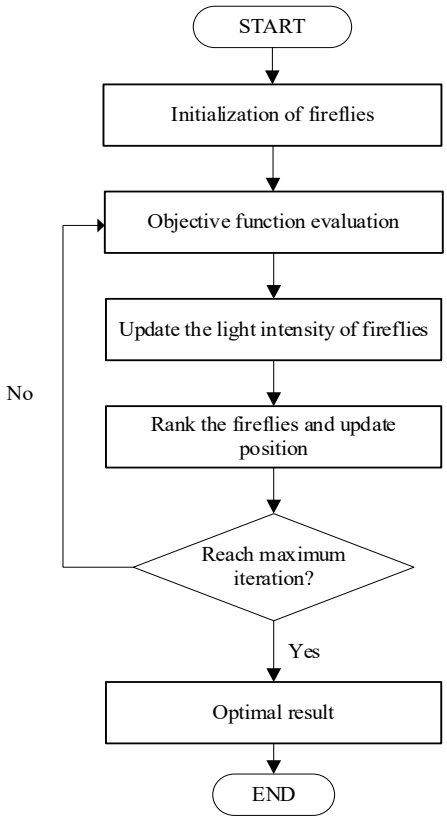


Figure 5: Flowchart of the Firefly Algorithm

In this study, the fuzzy system is used to generate fuzzy rule and membership function by using the FA. For the classification process, the fuzzy engine used was Sazonov Fuzzy Engine. It is a freely available fuzzy engine which is fully implemented in Java. It can be downloaded at http://people.clarkson.edu/~esazonov/.

3. Experimental Results

This section presents the analysis of the results obtained. Fuzzy modelling by using FA was tested with many parameters affecting the performance of the system . The effects of five parameter's performance which are population size, gamma probability, alpha0 probability, alphan probability, and the number of generations are shown in Table 2.

Table 2: Parameters setting

Parameter	Value
Population size	[10,40]
Gamma probability	[0.5,2.0]
Alpha0 probability	[0.1,0.4]
Alphan probability	[0.001,0.004]
Number of generations	[50,200]

As mentioned in most of the research papers, FA needs a small number of populations e.g. from 10 to 40 [14]. This is to make sure that the solution can be gained in a short time. For the gamma probability value, the current study stated that the ideal value for the gamma is 1.0 [15]. Next, the papers stated that 0.2 is the ideal value for the alpha0 parameter and 0.001 is the ideal value to be used for alphan probability [15]. Meanwhile, the number of generations tested was in a range of 50 to 100.

Table 3: The best parameters setting

Parameter	Value
Population size	20
Gamma probability	1.0
Alpha0 probability	0.2
Alphan probability	0.001
Number of generations	100

After the sensitivity analysis of every parameter in the algorithm has been performed, the best parameter setting can be found. The best parameter generated can be used to find the highest fitness value and is able to produce higher interpretability of the fuzzy model. Table 3 shows the best results for every parameter setting when being applied to both datasets.

Lastly, the accuracy of the results for every dataset was recorded and analyzed. The results were obtained after applying the best parameter value shown in Table 3. The first dataset shows that the highest accuracy, which is the best result, manages to reach up to 100% accuracy while the worst value is 97.72% and the average accuracy is 98.86%. Meanwhile, the second dataset indicates the value of 99.46% as the best accuracy value, 95.52% as the worst value, and 97.49% as average accuracy. To access the performance of the proposed study, the accuracy results for both datasets obtained were compared with other works. Table 4 gives the comparison of dataset 1 while Table 5 compared the results of dataset 2 with other works. Table 4 and Table 5 clearly show that the proposed method produces the best results compare to others.

Table 4: The comparison of results with other works for dataset 1

Work By	Result
Kaytan and Hanbay [16]	95.05%
Ubing et al.[17]	92.5%
Vrbančič, Fister, and Podgorelec [18]	94.4%
Mohd Foozy [19]	95.53%
This study	**98.86%**

Table 5: The comparison of results with other works for second dataset

Work By	Result
Mathew and Issac [20]	98.22%
Kawade[21]	98.34%
Raj et al. [22]	96.17%
Safie et al. [23]	97.13%
This study	**99.46%**

4. Conclusion

In this study, an improved method for fuzzy modelling was proposed and presented in detail. The method used FA for generating the fuzzy rule and membership function automatically for identifying the fuzzy parameter. To test the performance of the proposed method, two benchmark datasets were used. The results of the implementation of the method have been analyzed and it showed that the proposed method performed well compared to other works in terms of accuracy. To conclude, the implementation of FA for fuzzy modelling was able to produce good results.

Conflict of Interest

There is no conflict of interest declared by the authors.

Acknowledgment

Special appreciation to Universiti Malaysia Pahang for the sponsorship of this study by approve the Ministry of Higher Education (MOHE) for Fundamental Research Grant Scheme (FRGS) with Vot No. RDU190113.

References

[1] K. Leng, C. Lin, K. Wong, K. S. C. Yong, and W. King, "A new hybrid ensemble feature selection framework for machine learning-based phishing detection system," *Inf. Sci. (Ny).*, vol. 484, pp. 153–166, 2019. https://doi.org/10.1016/j.ins.2019.01.064

[2] M. M. Moreno-Fernández, F. Blanco, P. Garaizar, and H. Matute, "Fishing for phishers. Improving Internet users' sensitivity to visual deception cues to prevent electronic fraud," *Comput. Human Behav.*, vol. 69, pp. 421–436, Apr. 2017. https://doi.org/10.1007/978-3-319-09129-7_17

[3] R. B. Francisco, M. F. P. Costa, and A. M. A. C. Rocha, "Experiments with Firefly Algorithm," in *Computational Science and Its Applications -- ICCSA 2014*, 2014, pp. 227–236. https://doi.org/10.1007/978-3-319-09129-7_17

[4] M. Baykara and Z. Z. Gurel, "Detection of phishing attacks," in *2018 6th International Symposium on Digital Forensic and Security (ISDFS)*, 2018, pp. 1–5. https://doi.org/10.1109/ISDFS.2018.8355389

[5] S. Parekh, D. Parikh, S. Kotak, and S. Sankhe, "A New Method for Detection of Phishing Websites: URL Detection," in *Proceedings of the International Conference on Inventive Communication and Computational Technologies, ICICCT 2018*, 2018, no. Icicct, pp. 949–952. https://doi.org/10.1109/ICICCT.2018.8473085

[6] V. Patil, P. Thakkar, C. Shah, T. Bhat, and S. P. Godse, "Detection and Prevention of Phishing Websites Using Machine Learning Approach," in *Proceedings - 2018 4th International Conference on Computing, Communication Control and Automation, ICCUBEA 2018*, 2019, pp. 1–5. https://doi.org/10.1109/ICCUBEA.2018.8697412

[7] J. H. Li and S. De Wang, "PhishBox: An approach for phishing validation and detection," in *Proceedings - 2017 IEEE 15th International Conference on Dependable, Autonomic and Secure Computing, 2017 IEEE 15th International Conference on Pervasive Intelligence and Computing, 2017 IEEE 3rd International Conference on Big Data Intelligence and Compu*, 2018, vol. 2018-Janua, pp. 557–564. https://doi.org/10.1109/DASC-PICom-DataCom-CyberSciTec.2017.10

[8] L. MacHado and J. Gadge, "Phishing Sites Detection Based on C4.5 Decision Tree Algorithm," in *2017 International Conference on Computing, Communication, Control and Automation, ICCUBEA 2017*, 2018, pp. 1–5. https://doi.org/10.1109/ICCUBEA.2017.8463818

[9] T. Churi, P. Sawardekar, A. Pardeshi, and P. Vartak, "A Secured Methodology for Anti-Phishing," in *International Conference on Innovations in Information, Embedded and Communication Systems*, 2017, pp. 1–4. https://doi.org/10.1109/ICIIECS.2017.8276081

[10] W. Niu, X. Zhang, G. Yang, Z. Ma, and Z. Zhuo, "Phishing emails detection using CS-SVM," in *Proceedings - 15th IEEE International Symposium on Parallel and Distributed Processing with Applications and 16th IEEE International Conference on Ubiquitous Computing and Communications, ISPA/IUCC 2017*, 2018, pp. 1054–1059.

[11] L. S. Riza, C. Bergmeir, F. Herrera, and J. M. Benítez, "frbs : Fuzzy Rule-Based Systems for Classification and Regression in R," *J. Stat. Softw.*, vol. 65, no. 6, pp. 1–30, Jun. 2015. https://doi.org/10.18637/jss.v065.i06

[12] S. K. Rawat, "A Review on Spam Classification of Twitter Data Using Text Mining and Content Filtering," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 5, no. 6, pp. 485–488, 2015.

[13] A. J. Umbarkar, U. T. Balande, and P. D. Seth, "Performance evaluation of firefly algorithm with variation in sorting for non-linear benchmark problems," in *AIP Conference Proceedings*, 2017, vol. 1836. https://doi.org/10.1063/1.4981972

[14] X.-S. Yang, "Firefly algorithms for multimodal optimization," in *Stochastic Algorithms: Foundations And Applications*, 2009, p. 178. https://doi.org/10.1007/978-3-642-04944-6_14

[15] J. Kwiecień and B. Filipowicz, "Firefly algorithm in optimization of queueing systems," *Bull. Polish Acad. Sci. Tech. Sci.*, vol. 60, no. 2, pp. 363–368, 2012. https://doi.org/10.2478/v10175-012-0049-y

[16] M. Kaytan and D. Hanbay, "Effective Classification of Phishing Web Pages Based on New Rules by Using Extreme Learning Machines," *Anatol. J. Comput. Sci.*, vol. 2, no. 1, pp. 15–36, 2017.

[17] A. A. Ubing, S. Kamilia, A. Abdullah, N. Jhanjhi, and M. Supramaniam, "Phishing Website Detection: An Improved Accuracy through Feature Selection and Ensemble Learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 1, pp. 252–257, 2019. http://dx.doi.org/10.14569/IJACSA.2019.0100133

[18] G. Vrbančič, I. Fister, and V. Podgorelec, "Swarm Intelligence Approaches for Parameter Setting of Deep Learning Neural Network," in *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics - WIMS '18*, 2018, pp. 1–8. http://doi.acm.org/10.1145/3227609.3227655

[19] C. F. Mohd Foozy, R. Ahmad, M. A. Faizal Abdollah, and C. C. Wen, "A Comparative Study with RapidMiner and WEKA Tools over some Classification Techniques for SMS Spam," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 226, no. 1, 2017. https://doi.org/10.1088%2F1757-899x%2F226%2F1%2F012100

[20] K. Mathew and B. Issac, "Intelligent spam classification for mobile text message," *Proc. 2011 Int. Conf. Comput. Sci. Netw. Technol. ICCSNT 2011*, vol. 1, no. December, pp. 101–105, 2011.

[21] D. R. Kawade, "SMS Spam Classification using WEKA," *Int. J. Electron. Commun. Comput. Technol.*, vol. 5, no. ICICC, pp. 43–47, 2015. https://doi.org/10.1109/ICCSNT.2011.6181918

[22] H. Raj, Y. Weihong, S. K. Banbhrani, and S. P. Dino, "LSTM Based Short Message Service (SMS) Modeling for Spam Classification," in *Proceedings of the 2018 International Conference on Machine Learning Technologies*, 2018, pp. 76–80. http://doi.acm.org/10.1145/3231884.3231895

[23] W. N. H. W. Safie, N. N. A. Sjarif, N. F. M. Azmi, S. S. Yuhaniz, R. C. M. Yusof, and S. Yaacob, "SMS spam classification using Vector Space Model and Artificial Neural Network," *Int. J. Adv. Soft Comput. its Appl.*, vol. 10, no. 3, pp. 130–142, 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 1, 294-301 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Mitigation Methods of Voltage Disturbances and Harmonic Currents of PWM Two-Level Inverter Supplying The 6-Pulse Diode Rectifier

Jerzy Ryszard Szymanski*, Marta Zurek-Mortka

Kazimierz Pulaski University of Technology and Humanities in Radom, Faculty of Transport, Electrical Engineering and Computer Science, Radom, Poland

ARTICLE INFO

Article history:
Received: 14 October, 2019
Accepted: 23 January, 2020
Online: 10 February, 2020

Keywords:
Common-mode voltage
Current harmonics
Differential-mode voltage
PWM modulation
Three-phase voltage inverter
Two-level PWM inverter

ABSTRACT

In fast charging station for Battery Electric Vehicles (BEVs) the three-phase PWM (Pulse Width Modulation) inverters generate parasitic Differential-Mode Voltages (DMV) and Common-Mode Voltages (CMV). Parasitic voltages are a side effect of using the width modulation to shape the phase-to-phase inverter's voltage. In this article the authors present a mathematical description of the differential-mode (DM) and common-mode (CM) voltages and carry out their spectral analysis. Based on the spectral harmonics analysis, the authors present a method for filtration of harmonics of DM and CM voltages aimed at limiting the capacitance parasitic currents: due to DM voltage – phase-to-phase parasitic current and CM voltage – ground parasitic currents. A resonant filter was used to eliminate harmonics of the input rectifier of the DC/DC converter supplied from a voltage inverter. As a result of tests, the THD_i harmonic content factor was significantly reduced and ground leakage current was eliminated from the PE electric shock protection system.

1 Introduction

The DC/DC converter rectifiers for charging electrochemical energy stores used in charging stations of Battery Electric Vehicles (BEV) are non-sinusoidal receivers and generate current harmonics. When PWM inverters are the source of voltage supplying local AC networks, it is necessary to eliminate high-frequency differential harmonics of differential-mode DM voltage and common-mode CM voltage, which are disturbances resulting from the use of PWM modulation. It is also necessary to limit the harmonic content of low rows in the phase current, as they increase losses in the network. In the case of rectifier power supply from the transformer, the phase and thus phase-to-phase voltages are deformed, transformer voltage deformations, determined by the THD_u harmonic content ratio, must not exceed of the permissible values [1],[2].

Scientific publications present solutions using transformers to galvanically separate the inverter, which converts the energy of the photovoltaic source from the input rectifier of the AC/DC/DC converter of the EV charger [3]. An interesting solution seems to be the use of a three-winding Yyd transformer in the 12-pulse diode rectifier of the AC/DC/DC charger [4]. One of the authors' research goals is to demonstrate the possibility of obtaining a level of harmonic content in the inverter current of a value close to when using a 12-pulse rectifier with a Yyd transformer.

This paper consist of theoretical part where the PWM inverter and the 6-pulse diode rectifier are analyses. It also shows the concept of powering from renewable energy sources dedicated for fast charging stations. Next part describes the solution of limiting disturbances of voltage inverters in hybrid system.Thanks to use of novel configuration of filters against voltage disturbances it is possible to mitigate the influence of rectifier on power supply system in charging station. The last part shows the results of simulations of rectifier's harmonics compensator and laboratory experiments to determine the conditions for cooperation of PWM converter with DC/DC converter using the CM and DM filters.

This article is an extension of the issues presented in the reference [5] and also is a continuation of the research presented there.

2 Hybrid Powering of Power Electronic Converters of EV Charging Station

In order to assure of high power and using of "clean" electricity, the electric vehicles charging station should be powered from hybrid system with an energy storage. Figure 1 shows an exemplary implementation of a hybrid power supply system for electric vehicles charging station with using the renewable energy sources.

In a hybrid power system a few energy sources is used. The loads are supplying from local AC or DC voltage supply network.

*Corresponding Author: Jerzy Ryszard Szymanski, phone: +48 601 292 535, e-mail: j.szymanski@uthrad.pl

In the hybrid network shown in Figure 1, the local DC voltage line is connected to renewable energy sources (RES) (e.g. wind farms or solar power plants) and energy storages, like batteries, fuel cells or pumped storages. The AC line is powering also DC receivers which are DC fast charging stations for electric vehicles. Solutions are known in which electric vehicles are connected charging station supplies energy to the power grid, here two-way DC/DC converters are used as chargers for electric vehicles EV. The energy of electrochemical batteries can also be sent to an industrial power supply network. At the same time, a power system that is synchronized with the local hybrid power network can complement local sources of "clean" energy from periods of reduced efficiency (e.g. on cloudy days for solar sources). In periods of increased efficiency of local renewable energy sources, excess energy can improve the efficiency of the power system. In such systems, bidirectional energy converters must be used [5], [6].

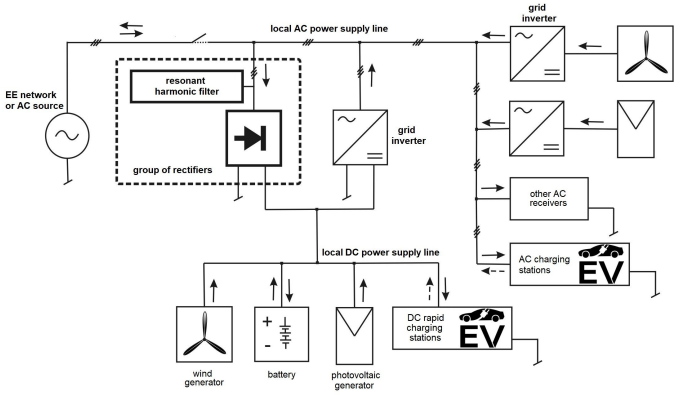


Figure 1: Local hybrid power supply system for EV charging stations with renewable energy sources (wind and PV)

Figure 2 shows the EV fast battery charging station, which is powered from a two-level voltage inverter or from the power grid. The passive parallel harmonic current filter is connected to the AC line with increased current consumption by the diode rectifier. This allows you to control the amount of capacitive reactive power supplied to the AC network by this filter. The passive filter used limits the level of harmonics in the THD_i range from 5% to 10% [5]. Powering the EV fast charging station from the low-voltage industrial network leads to distortion of phase and phase-to-phase voltages in this network. Voltage distortions increase due to the power of the battery charging station, in particular at similar capacities of MV / LV transformers in the power system. Such voltage distortions cause disturbances causing incorrect operation of other receivers connected to the power supply network, e.g. contactors, PLC controllers and other [7]. Limiting the current harmonic content is necessary today, because the number of diode rectifiers and their power is constantly increasing.

For the purposes of this article, it was assumed that the MV/LV transformer of the power system supplies, apart from the fast EV battery charging station, also other electric receivers. In places where travelers service on expressways and highways, where fast EV battery charging stations are located, transformers usually power other receivers, such as pumps, lighting, air conditioners. It was assumed that the fast charging station does not have its own dedicated MV/LV transformer, which is why the use of effective limiting harmonics of phase currents is very important.

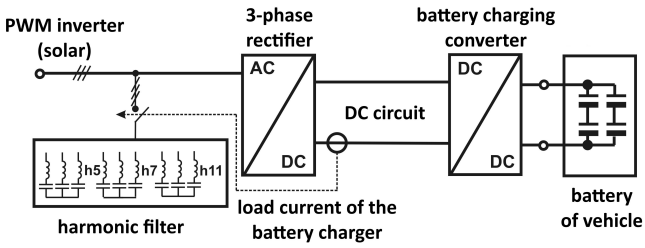


Figure 2: Block diagram of an EV battery charger with a passive harmonic current filter to provide sinusoidal current in the local power supply network

3 Mathematical Description of a Linear Three-Phase Inverter Model

The transformation of DC voltage into single-phase and three-phase AC voltage, with the use of the power-electronic PWM inverter, has side effects such as the CM voltages [8], [9]. In this chapter, the mathematical relations describing the voltage of the common mode disturbances were derived bases on the description of the three-phase linear model of the PWM inverter. The three-phase two-level voltage inverter is shown with a symmetrical load in Figure 3. It consists of three identical branches of half-bridges in each phase and this three-phase system is symmetrical for individual harmonics. In modeling and analyzing the inverter, a three-phase notation is used [10], [11].

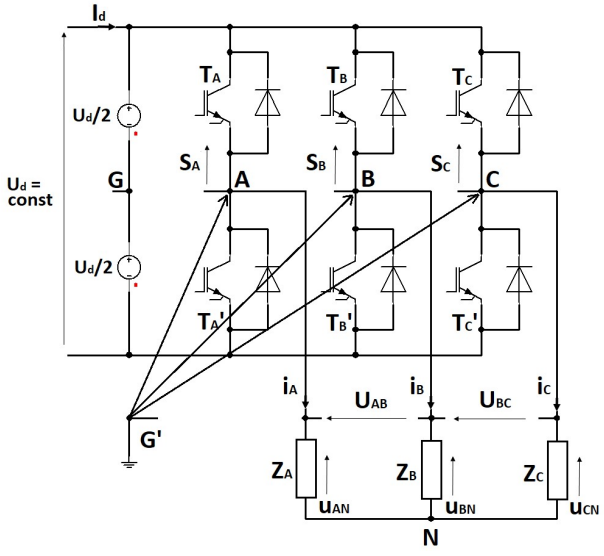


Figure 3: Basic scheme of three-phase two-level inverter with symmetrical load

On the basis of the three-phase inverter description, a mathematical relationship is determined, which describe the CM voltage of the inverter with the grounded negative DC voltage bus and without

galvanic ground of the power-electronic elements (such as an industrial inverters and auxiliary railway equipment inverters) [10], [12].

The output phase-to-phase voltages of the inverter are described as follow 1:

$$\begin{aligned} u_{AB} &= u_{AG} + u_{BG} \\ u_{BC} &= u_{BG} + u_{CG} \\ u_{CA} &= u_{CG} + u_{AG} \end{aligned} \tag{1}$$

where: u_{AG}, u_{BG}, u_{CG} are instantaneous voltages, which satisfy the condition 2:

$$u_{AB} + u_{BC} + u_{CA} = 0 \tag{2}$$

If the load neutral point N is not connected to the reference point G on the DC voltage side (Figure 3), then the phase voltages can be expressed as 3:

$$\begin{aligned} u_{AN} &= \tfrac{1}{3}(2u_{AG} - u_{BG} - u_{CG}) \\ u_{BN} &= \tfrac{1}{3}(-u_{AG} + 2u_{BG} - u_{CG}) \\ u_{CN} &= \tfrac{1}{3}(-u_{AG} - u_{BG} + 2u_{CG}) \end{aligned} \tag{3}$$

By describing the switching states of inverter S_A, S_B, S_C with 0 and 1 (e.g. S_A=1, indicates the state in which the transistor in phase A is switched on, and S_A=0, when the lower transistor in phase A is switched off), the output phase-to-phase voltages can be presented as follows 4:

$$\begin{aligned} u_{AB} &= U_d(S_A - S_B) \\ u_{BC} &= U_d(S_B - S_C) \\ u_{CA} &= U_d(S_C - S_A) \end{aligned} \tag{4}$$

Similarly for a symmetrical load the phase voltages are expressed by 5:

$$\begin{aligned} u_{AN} &= \tfrac{U_d}{3}(2S_A - S_B - S_C) \\ u_{BN} &= \tfrac{U_d}{3}(-S_A + 2S_B - S_C) \\ u_{CN} &= \tfrac{U_d}{3}(-S_A - S_B + 2S_C) \end{aligned} \tag{5}$$

If the G point is grounded in the Figure 3, then the u_{NG} voltage is determined by CM voltage of inverter 6:

$$\begin{aligned} u_{AG} &= u_{AN} + u_{NG} \\ u_{BG} &= u_{BN} + u_{NG} \\ u_{CG} &= u_{CN} + u_{NG} \end{aligned} \tag{6}$$

The sum of all voltages from equations 5 is 6:

$$u_{AG} + u_{BG} + u_{CG} = u_{AN} + u_{BN} + u_{CN} + 3u_{NG} \tag{7}$$

Since $u_{AG} + u_{BG} + u_{CG} = 0$, the CM voltage u_{NG} of inverter is determined by 8:

$$u_{NM} = \frac{u_{AN} + u_{BN} + u_{CN}}{3} \tag{8}$$

The relation 8 indicates that the CM voltage of inverter is equal to the voltage with respect to another ground point, e.g. G' (Figure 3), which is not galvanically connected to the inverter's electrical circuit (e.g. relative to the protective PE ground of the drive system). However, if the ground potential point G' is shorted with the negative DC voltage bus, then according to the Figure 3 the voltage $u_{GG'} = U_d/2$ and then the following applies 9 [11]:

$$u_{NG'} = u_{NG} + u_{MG'} = u_{NG} + \frac{U_d}{2} \tag{9}$$

Taking into account 9 and the voltages from the equation 6, a dependency considering the sum is obtained 10:

$$u_{NG'} = \frac{u_{AN} + u_{BN} + u_{CN}}{3} + \frac{1}{2}U_d \tag{10}$$

From 10 a conclusion can be drawn that the grounding of the negative DC voltage bus of the inverter does not eliminate the inverter's common-mode voltage $u_{NG'}$, but instead only increases its level relative to the ground. The grounding of the negative DC voltage bus of the inverter causes the ground disturbances current caused by this voltage, directly flow into the voltage inverter. It is a property caused a higher impedance in industrial frequency converters, because in this type of inverter's there is no galvanic connection between the negative DC voltage bus and ground [11].

4 Limiting Disturbances of High-Frequency Voltage Inverters in a Hybrid Grid

The DM and CM voltage of the PWM inverter are its incherent feature [12], [13]. The filtration of these undesired voltages eliminates a number of undesirable phenomena that occur when high-frequency capacitive leakage currents occur: phase-to-phase current from DM voltage and ground current from CM voltage. The use of significant inductance in the DM voltage filter, which simultaneously impedance separates the voltage inverter from the resonant current harmonic filter, simultaneously strengthens the level of current harmonic filtration.

The use of DM and CM voltage filtration is of fundamental importance for limiting the spread of high-frequency disturbances in the hybrid power network [12], [14]. High-frequency ground leakage capacitance currents usually flow through residual current circuit breakers without causing them to respond to an exceeded ground fault current, resulting in a risk of electric shock [15], [16]. Non-sinoidal phase voltages, e.g. in the form of a symmetrical rectangular waveform, will always generate a CM rectangular voltage. Rectangular phase voltage with phase shifts close to zero arise with the sinusoidal PWM modulation of the inverter and the modulation depth factor M is close to zero.

In the inverter model from 5 supplying the diode rectifier from Figure 10, a parallel resonance filter was used to limit harmonics in phase currents. Inductance (L3, L4, L5) of the DM voltage filter increases the impedance between the inverter and the resonant current filter, thereby increasing its efficiency [5], [13].

Suppression of the CM voltage harmonics of the PWM inverter is possible by forcing the capacitive leakage current CM as shown in Figure 5. When the AM1 ammeter current is zero, only the harmonics of the DM voltage are suppressed and the inverter phase-to-phase voltages are sinusoidal.

When the i_0 = AM1 current is different from zero then additionally the phase voltages are also sinusoidal and the CM voltage can be described by the formula 11:

$$u_{CM} = \frac{1}{3}\frac{1}{C_f}\int i_0 dt \tag{11}$$

where: i_0 = current at AM1, C_f - filter phase capacitance (Figure 5: C6, C7, C8 and if C5 $\gg C_f$ the $u_{CM} = u_{C5}$).

Formula 11 shows that with sufficiently large phase capacities C_f of the DM voltage filter it is possible to obtain phase voltages in the form of pure sinusoidal waveforms. The voltages obtained from the inverter (e.g. photovoltaic) supply the premises with a three-phase AC voltage line to which single- and three-phase receivers can be connected. Therefore it should be free of harmonics deforming phase voltages. In the present case, the AC line is loaded with the diode rectifier of the EV battery charger (Figure 10). The rectifier battery load models the load while charging the battery storage (e.g. DC/DC loaded converter). When galvanic connection of the inverter power outputs with the rectifier power inputs, a sufficiently large inductance value in the current path must be ensured. In the simulation tests, the inverter sinusoidal PWM modulation with a carrier frequency of 6kHz is used (modulation depth factor M = 1).

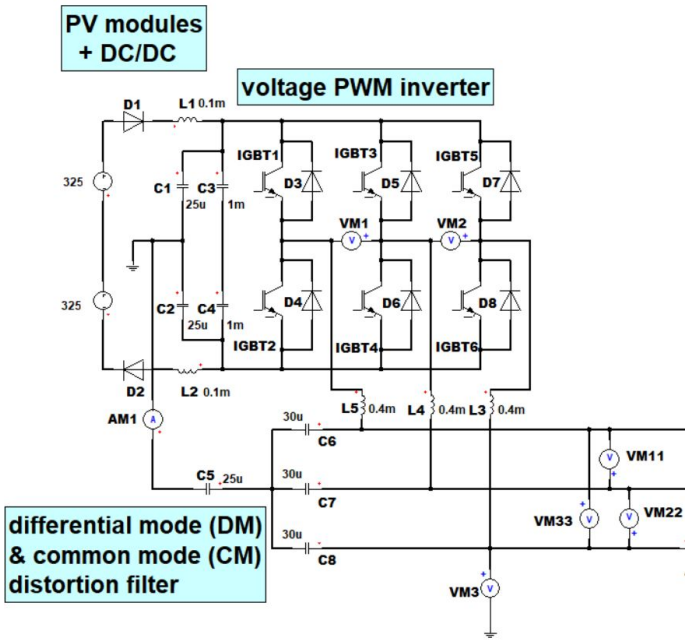


Figure 5: Two-level PWM inverter with DM and CM voltages filters with local AC line to supply of diode EV charger rectifier

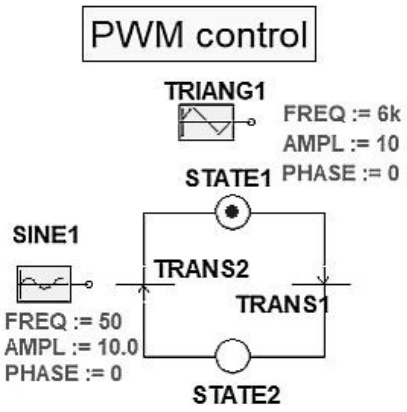


Figure 4: Event graph for PWM control of a three-phase inverter for one phase

The inverter PWM control in the form of an event graph is shown in the Figure 4. The carrier section is a symmetrical triangle with a frequency of 6kHz. Modulating waveforms are 50 Hz sine waves. In each phase they are phase shifted by an angle of 120°. For conducted tests of current harmonics produced in the rectifier and CM and DM inverter voltages, sinusoidal PWM modulation is used.

In the inactive phase of the inverter (i.e. when energy is not supplied by the inverter), the energy accumulated in the current path inductances maintains a sufficiently large rectifier supply current. The LC inductance of the DM disturbance filter accumulates energy (W_L=0.5LI^2), which supports the flow of phase currents to the rectifier when PWM modulation blocks the flow of currents through the inverter IGBTs.

The model of inverter with PWM sinusoidal modulation to feed diode rectifier in electric vehicles charging station is shown in the Figure 5. The reference [17] proposes other system structure with use the isolation transformer or boost converter [18] with inverter [1]. The inverter is powered by a solar power plant using PV modules (solar energy). The inherent DM and CM voltages of the PWM inverter are minimized with the DM filter (phase-to-phase harmonics) and CM filter (phase voltage harmonics after filtering the DM harmonics first) [6].

The PWM inverter model with DM and CM voltage filters used in simulation tests is shown in Figure 5. It enables simulation tests of the hybrid power system [5]. Ansys Simplorer software was used to build the model and perform simulation tests.

The phase-to-phase waveforms of the inverter obtained in simulation tests indicate that the parallel resonant harmonic filter in phase currents strengthens the inverter's DM voltage filtration efficiency. Comparing the phase-to-phase voltages of the inverter shown in Figures 6 and 7, a nearly 10% reduction in the THD_u harmonic content factor is visible when the current harmonic filter is on.

The voltmeters VM1 and VM11 used in the voltage inverter model from Figure 5 indicate the phase-to-phase voltages before and after the DM voltage filter, respectively. Figure 6 shows voltages when switched off and figure 7 when the resonant current harmonic filter is switched on.

Occurring surges in the phase-to-phase voltage lead to numerous undesirable phenomena, including resonance phenomena. Suppression of higher harmonics at phase-to-phase voltage is fundamental when limiting the high frequency disturbances spread in the local AC network. The beneficial effect of the applied harmonic phase current filter on increasing the harmonic suppression of the DM voltage in the phase-to-phase voltages of the inverter is shown in Figure 7.

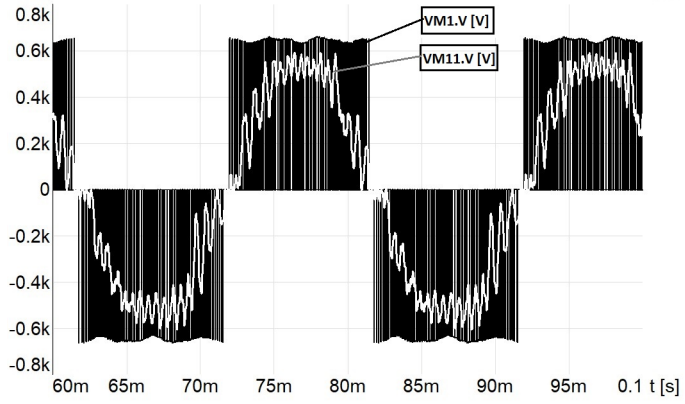


Figure 6: Phase-to-phase voltage of loaded PWM inverter (VM1) before DM and CM filters and after DM and CM filters (VM11 - THD_{1u}=17.55%) when resonance harmonic filter is **off**

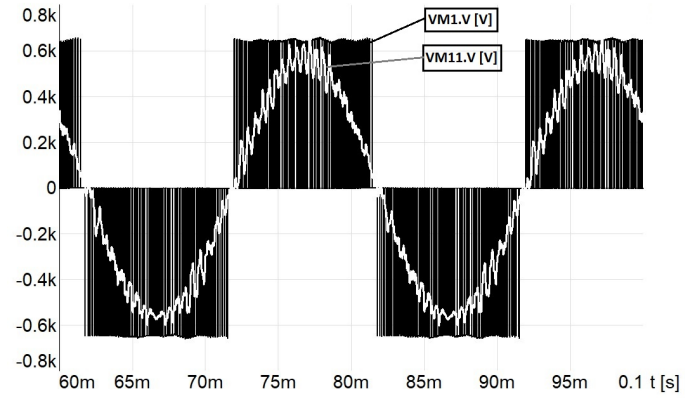


Figure 7: Phase-to-phase voltage of of loaded PWM inverter (VM1) before DM and CM filters and after DM and CM filters (VM11 - THD_{1u}=7.65%) when resonance harmonic filter is **on**

5 Model and Research of a Current Resonant Filter in a Three Phase Rectifier System Powered by a Voltage Inverter

Fast EV charging stations can be DC or AC powered. DC supply from renewable energy generators or electrochemical reservoirs (e.g. lithium-ion storage) relieves the local AC network. Power supply for the input rectifiers of EV battery chargers (DC/DC converters) requires harmonic current filtering, e.g. by using 5, 7 and 11 harmonic passive parallel filters. The harmonic filter is attached to a three-phase AC line depending on the rectifier's current rating. At low current diode rectifiers, the filter is disconnected from the AC line to limit the reactive capacitive power supplied to the power system (Figure 2).

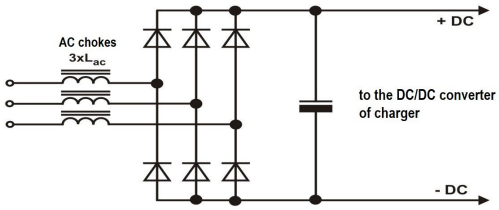


Figure 8: Chokes at the diode rectifier inputs to limit the steepness of the capacitor bank charging current pulses - AC chokes

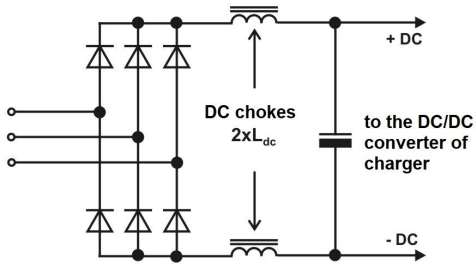


Figure 9: Chokes on the DC voltage side of the diode rectifier to limit the steepness of the current pulses of the capacitor bank charging - DC chokes

Three-phase diode rectifiers are equipped with AC chokes - Figure 8 or DC - Figure 9, or a combination thereof. These are usually chokes that cause a phase voltage drop from 3% to 4% at nominal current, which results in providing THD_i harmonic content ratio below 45% [1]. AC chokes protect the AC line against switching commutation of current in rectifying diodes, but they cause undesirable reduction of the rectified voltage value. DC chokes do not lower the straightened tension. If the source of AC voltage supplying the rectifier in the hybrid power supply network is a voltage inverter, then when determining the value of the AC or DC choke inductance of the rectifier, the inductance of the inverter's DM voltage filter should be taken into account. Usually the AC and DC chokes have similar inductance values, which decrease proportionally with increasing power of the rectifiers [5]. DC reactors are placed symmetrically in the branches $+U_{DC}$ and $-U_{DC}$ to ensure a symmetrical CM voltage of the rectifier relative to the ground potential, when it is fed from a power network with a TN network system.

The values of the capacity of the batteries and chokes used in the diode rectifiers for 50kW and 500kW EV converters are summarized in Table 1. The proportional relationship between the capacity of the EV rectifier batteries (figures 8 and 9) and its power is visible here, and the inductance of the chokes is inversely proportional.

Table 1: Basic parameters of EV rectifiers with power 50kW and 500kW

$P_r[kW]$	50	500
$C_{DC}[mF]$	10	100
$L_0(4\%) = L_{DC}[mH] = L_{AC}[mH]$	0.3	0.03

Using the L_0 AC or DC chokes in the rectifier with a C_{DC} capacitor battery EV charger with values from the table 1 reduces the harmonic content THD_i in the phase current of the AC line to a value of approx. 40%-45% [1]. Further reduction of the harmonic content coefficient, to the level of nearly 12% (Figure 12), was obtained by using a three-branch passive resonance filter (for 5th, 7th,

and 11th harmonic). Obtaining phase sinusoidal currents increases the efficiency of the EV charger rectifier supply system [6]. The AC and DC inductance of the chokes presented in table 1 take into account the inverter DM voltage filter phase inductance connected in series or the transformer inductance when the EV charger is powering from the power grid.

6 Model and Simulation Tests of The Rectifier's Harmonic Current Passive Filter

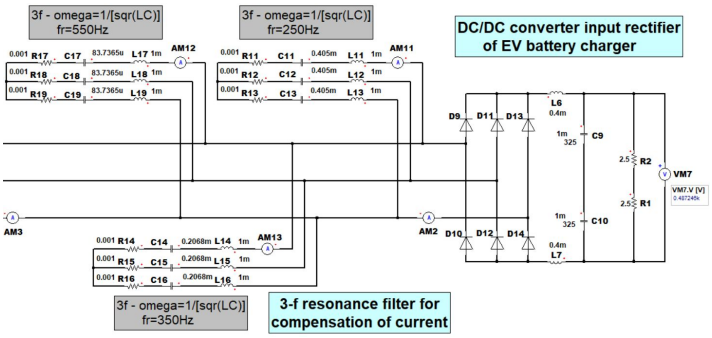


Figure 10: Model of 6-diode rectifier with parallel resonant filters for 5th, 7th and 11th harmonics

Diode rectifiers are commonly used to convert AC voltage to DC. In the DC/DC converter, the voltage is processed in such a way as to maintain a constant allowable current for charging the EV battery (Figure 1). In the rectifier model shown in Figure 10, the individual components mean: bank of capacitors (C9, C10), rectifier resistance load (R1 + R2), harmonic resonant filters for 5th, 7th and 11th harmonics.

Fast Fourier transform was used to set the coefficient of the THD_i phase harmonics. The results of simulation tests are shown in Figures 11 and 12.

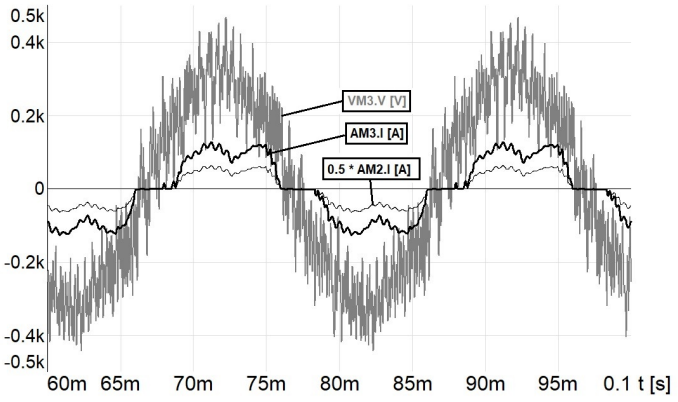


Figure 11: Phase voltage and currents waveforms on AC local network with the PWM inverter and diode rectifie: VM3 – phase voltage before rectifier, AM2 (THD_i=45,4%) current before the harmonic filter, AM3 - current after the harmonic filter when filter is off (Figure 10)

The phase current harmonics in Figures 11 and 12 are effectively filtered by means of a passive resonant filter. The high efficiency of the resonant filter is obtained by using the inductance of the inverter's DM voltage filter (Figure 5). In simulation studies, effective harmonic filtration has been demonstrated in the power supply system of a voltage inverter - diode rectifier, without the use of a separating transformer between these converters.

Figures 11 and 12 also show the VM3 phase voltage of PWM inverter loaded with active power of 50kW. The sinusoidal shape of the VM3 phase voltage indicates the effective operation of the CM voltage filter on PWM inverter outputs (Figure 5).

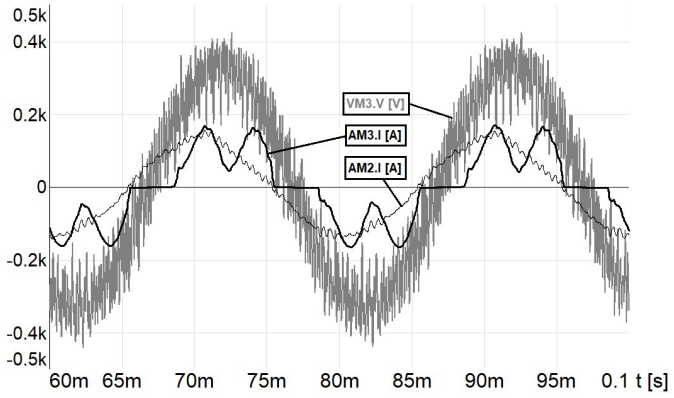


Figure 12: Phase voltage and currents waveforms on AC local network with the PWM inverter and diode rectifier: VM3 – phase voltage before rectifier, AM2 - current before the harmonic filter, AM3 (THD_i=12.21%) - current after the harmonic filter when filter is on (Figure 10)

7 Laboratory Experiments Results

Laboratory stand to testing cooperation of PWM volatge inverter (1) with rectifier (7) of drive inverter loaded with motor (11) is presented on the Figure 13.

The stand was built to examine the cooperation between a voltage inverter (for example, as a voltage source of a photovoltaic power plant) and a diode rectifier loaded with capacitive battery (for example, as an input stage of the DC/DC converter of an electrochemical energy storage charger). This solution does not use a separating transformer between the inverter and the rectifier, because it is an expensive solution, with large dimensions and weight. The authors are looking for an alternative (low cost) solution without using transformer separation between the inverter and the rectifier.

The main purpose of the research is to determine the conditions for cooperation of a PWM voltage converter generating three-phase alternating voltage with a DC/DC converter for charging EV batteries as shown in the Figure 1.

The laboratory stand, depicted on the Figure 13, for AC nominal line currents 3x10A/50Hz and phase-phase voltages 3x400V/50Hz contains:

- LC filter (3xL=4mH, 3xC=3μF) at the inverter output (differential-mode voltage filter) (4), which limits high frequency disturbances in phase-phase voltages of inverter (1),
- radio frequency interferences filter RFI (2) of the inverter (1),
- branch of CM voltage filter: 2xC=2x15μF, L=1mH (5),(6),

- diode rectifer of the voltage inverter (7),
- excess energy dissipation resistor (8),
- load system (11),(12) of the voltage inverter (7).

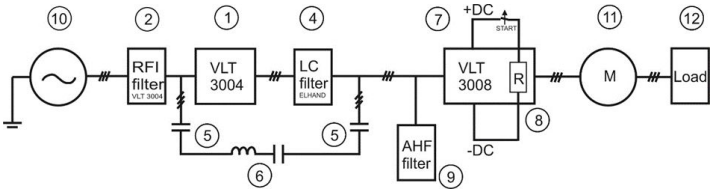


Figure 13: Block diagram of the laboratory stand for testing the cooperation of the renewable energy voltage inverter with the diode rectifier of the vehicle battery charging system

Figure 14 shows the measured line current and phase voltage on the power supply of the loaded diode rectifier of AC/DC/AC converter with 10A current. A significant deformation of the line current in the measured phase is visible as a result of charging the rectifer's (AC/DC converter) capacitor banks (THD_i approx. 45%).

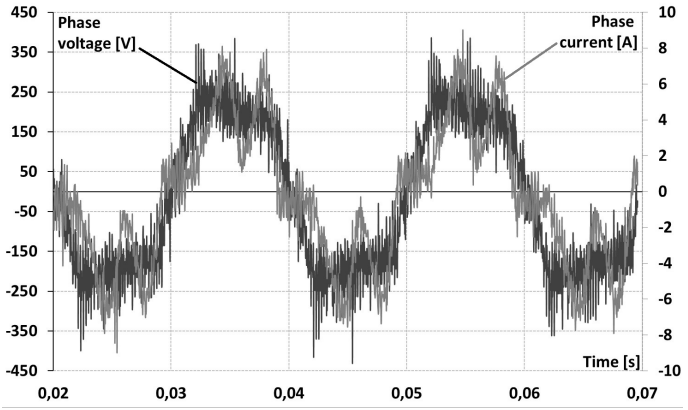


Figure 14: Phase voltage and phase current after the LC filter (4) at the test stand shown in the Figure 13 with the passive harmonic filter AHF [order: 5, 7, 11] switched **off** (9)

The phase voltage has a limited harmonic content of high orders, which confirms the correct operation of the CM filter (Figure 13). Previous simulation tests of the model of the local power supply system for vehicle battery charging stations are consistent with the experimental results (Figure 11) The capacitive reactive power consumption is not significant here,which results from a negligible phase shift of the first harmonic of the current and phase voltage.

The increased consumption of capacitive reactive power is observed after connecting the AHF passive harmonic filter (orders: 5, 7, 11), also under nominal load conditions, as shown in the oscillogram from the Figure 15 and the values are $\cos\phi_1$=0.89, ϕ_1=27°.

The oscillogram from the Figure 15 confirms the effective filtration of low order harmonics through the attached parallel resonant hamonic filter. The high efficiency of the LC passive filter was obtained by using nearly 3% current path inductance (i.e. 3% phase voltage drop on the inductance at nominal current), such as rectifier DC inductance (Figure 9) and DM disturbance filter inductance (Figure 5). Oscillogram from the Figure 15 also shows that the use of a resonant parallel harmonic filter has enabled the reduction of phase current harmonics to THD_i = 12.2% from above 45%. The experimental result confirmed the previously obtained simulation test result of the model shown in the Figure 12.

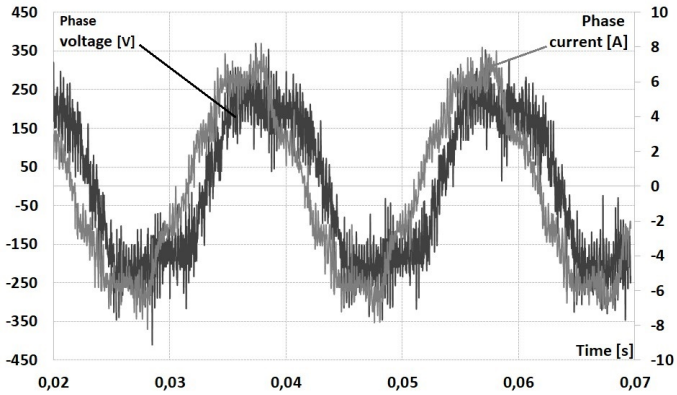


Figure 15: Phase voltage and phase current after the LC filter (4) at the test stand shown in the Figure 13 with the AHF passive harmonic filter [order: 5, 7, 11] switched **on** (9)

8 Conclusions

The presented system from transformerless supply of a 6-diode rectifier from a two-level PWM inverter is a low cost alternative to supplying the rectifier using a transformer. The tested power supply system can be used in fast EV charging stations. The system enables building a scalable EV charging station composed of modules, e.g. 50kW each.

In the tested system, the current efficiency of the inverter is comparable to the load power of the rectifier.

The applied CM and DM voltage filtering of the inverter made it possible to obtain practically sinusoidal phase and phase-to-phase voltages in the local three-phase AC network.

By using a passive resonant current harmonic filter, practically sinusoidal phase currents in the AC network were obtained at high loads of the diode rectifier, which increases the efficiency of the tested power supply system.

The applied DM differential voltage filters and CM common voltage provide a sinusoidal shape of phase and phase-to-phase voltage in the three-phase local AC network. The inductance of the 3xL DM filter also allows the accumulation of energy necessary to maintain the rectifier current continuity. In the inactive phase of the inverter (operation with passive control vectors: 000 or 111), the energy accumulated in the DM filter inductances is transferred to the rectifier.

To obtain sinusoidal phase and phase-to-phase voltages in the local hybrid three-phase voltage network, it is necessary to use DM and CM voltage filters of the inverter. The inductance of the DM voltage filter enhances the efficiency of the resonant harmonic filter of phase currents. If the CM filter is using, an additional energy storage to limit the voltage rise above the permissible value in the inverter's DC circuit is need to consider. In case the PWM voltage inverters feed the rectifier circuits, it is necessary to use passive or active harmonic filter. The use of a passive harmonic current filter

between the inverter and the rectifier has made it possible to reduce the THD_i of phase currents to a level similar to that of 12-pulse rectifiers with a transformer.

Simulation and experimental tests performed of the hybrid network model using the PWM inverter, filters and the rectifier showed the possibility of obtaining three-phase voltage source of PWM inverters with limited harmonic content.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgments In the part of the article related to the modelling of the electric circuit, the ANSYS Software national scientific software license has been used, which was funded by a computational grant obtained by Kazimierz Pulaski University of Technology and Humanities in Radom, Poland.

References

[1] Hansen S., Asiminoaei L., Blaabjerg F., Simple and Advanced Methods for Calculating Six-Pulse Diode Rectifier Line-Side Harmonics, 38th IAS Annual Meeting. Conference Record of the Industry Applications, vol. 3, 2003, pp. 2056 – 2062, DOI:10.1109/IAS.2003.1257849.

[2] Szymanski J., Common and differential mode voltage in frequency conventer drive systems. ElektroInfo, vol. 11, 2013. (article in Polish with an abstract in English).

[3] Li X., Xing X., Zhang Ch., Chen A., Qin Ch., Zhang G., Simultaneous Common-Mode Resonance Circulating Current and Leakage Current Suppression for Transformerless Three-Level T-Type PV Inverter System, IEEE Transactions on Industrial Electronics, Volume: 66 , Issue: 6 , June 2019, DOI: 10.1109/TIE.2018.2860555.

[4] Trivedi N., Gujar N., Sarkar S., Pundir S.P.S., Different fast charging methods and topologies for EV charging, 2018 IEEMA Engineer Infinite Conference, DOI: 10.1109/ETECHNXT.2018.8385313

[5] Szymanski J., Zurek-Mortka M., Harmonic Resonant Filters of Quick Battery Charging Station of Motor Vehicles, 23rd International Conference Electronics, IEEE, 2019, DOI: 10.1109/ELECTRONICS.2019.8765691

[6] Zurek-Mortka M., Szymanski J.: Reactive power compensation in a harmonic environment. elektro.info, vol. 10, 2018. (article in Polish with an abstract in English).

[7] Pedzisz K. Capacitor banks measuring analysis and operating condition assessment in industrial environment, Journal Electrical engineering and electronics, vol. 27, p.2, 2008.

[8] Kempski A., Smoleński R., Decomposition of EMI Noise Into Common and Directional Modes in PWM Inverter Drive System. Electrical Power Quality and Utilization, 12 (no. 1), 2006.

[9] Pairodamonchai P., Sangwongwanich S., Exact Common-Mode And Directional-Mode Equivalent Circuits of In-verters in Motor Drive Systems Taking Into Account Input Rectifiers. IEEE PEDS'11, Singapore, 5-8 December, 2011.

[10] Ambrožič, V., Zajec P., Električni Servo Pogoni, Slovensko združenje elektroenergetikov CIGRE-CIRED, Ljubljana 2016.

[11] Szymanski J., The Voltage of Common-Mode Disturbances in Frequency Converters in Railway Systems. Monography, UTH Radom, 2015.

[12] Tallam R. M., Skibinski G. L., Shudarek T. A., Lukaszewski R. A. Common-mode voltage reduction PWM algorithm for AC Drives, IEEE Transactions on Industry Applications, vol. 46, no. 5, pp. 1959–1969, 2010. DOI: 10.1109/TIA.2010.2057396.

[13] Takahashi S., Ogasawara S., Takemoto M., Orikawa K., Tamate M., Common-Mode Voltage Attenuation of an Active Common-Mode Filter in a Motor Drive System Fed by a PWM Inverter, IEEE Transactions on Industry Applications, pp.1-1, 2019. DOI: 10.1109/ICEMS.2017.8056376.

[14] Danfoss, Output filters design guide - MG90N402, 2010.

[15] Czapp S. Operation of residual current devices at an increased frequency of differential current. ElektroInfo, vol.10, 2008. (article in Polish with an abstract in English).

[16] Charoy A., Dunand Ph., Bearing currents induced by a power drive. Automotive Power Electronics, 26 - 27 Sept., Paris, pp.1-7, 2007.

[17] Park J.K., Kim Y.H., Jung Y.Ch., Won Ch.Y., A novel control strategy of an active clamped flyback inverter with synchronous rectifier for a photovoltaic AC module system, 8th International Conference on Power Electronics – ECCE Asia, 2011. DOI: 10.1109/ICPE.2011.5944557.

[18] Genikomaskis K.N., Gutierrez I.A., Thomas D., Ioakimidis C.S., Simulation and design of fast-charging infrastructure for a university-based e-Carshring System, IEEE Transactions of Intelligent Transportation Systems, vol.19, no.9, 2018. pp. 2923–2932. DOI: 10.1109/TITS.2017.2767779

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 5, 294-299 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Spectrum Requirements for Control and Non-payload Communication of Maritime Autonomous Surface Ship

Ho Namgung*

Graduate School of Maritime Transportation System, Mokpo National Maritime University, Mokpo, KS005, Republic of Korea

ARTICLE INFO

Article history:
Received: 30 July, 2019
Accepted: 06 September, 2019
Online: 15 October, 2019

Keywords:
MASS
CNPC
LTE-M
ITU-R

ABSTRACT

This paper proposed a frequency bandwidth required for the Control and Non-Payload Communications (CNPC) in autonomous operations of the Maritime Autonomous Surface Ship (MASS). The issue was to seek a radio spectrum to meet future demand for safe operations of autonomous ship at sea. As described in Document 5B/841 of ITU-R SG, safe navigation of autonomous ships relies on communications for command and control. Based on the operations required for the safe navigation of the MASS, the total bit rates were calculated. The spectrum requirements were obtained for command and control communications in operating the MASS safely. As a result, in terms of bandwidth, transmit rate, and transmit distance, state of the art satellite communication system was suitable for a safe operation of the MASS.

1. Introduction

Recently the Maritime Autonomous Surface Ship (MASS) has received much attentions from maritime institutions and industries. The Maritime Safety Committee (MSC) of the International Maritime Organization (IMO) agreed in its 99th session that the Committee provides an output on "Regulatory scoping exercise for the use of MASS" with a target completion year of 2020 [1]. In 2015, ITU-R Study Groups Working Party 5B has requested member states to study and provide a basis for considering regulatory, technical and operational conditions, for the Control and Non-Payload Communications (CNPC) link between satellite services, fixed service, mobile service, and the Unmanned Maritime Systems (UMS) in maritime environments [2]. The society of maritime industries on behalf of United Kingdom has released the MASS Code of Practice Version 1.0 which generally describes all the operations of the MASS including communications systems. As described in the 95th session of MSC [3], significant academic research and commercial research and development (R&D) have undertaken on technological aspect of the autonomous navigation and collision avoidance systems. In particular, the Maritime Unmanned Navigation through Intelligence Network (MUNIN) project proposed internal & external communication systems based on state-of-the-art satellite communication systems and communication requirements [4],[5]. In Finland, data requirements for wireless transmission regarding the environmental sensors and remote maintenance as well as remote control needs were identified for connectivity challenges of autonomous ships in different environments, including ports, deep sea, and Arctic regions [6].

However, there are issues as follows: (i) did not apply a concrete methodology for securing communication channels which are dedicated for safe operation of the MASS via the CNPC; and (ii) only used state-of-the-art satellite communication system based on communication requirements without calculation of spectrum requirements.

Therefore, in this paper, the spectrum requirements among the available maritime frequency band was calculated to secure appropriate communication channel for safe operation of the MASS by analyzing and utilizing the density of ship traffic. Section 2 described methodology for spectrum requirements, and Section 3 calculated the spectrum requirements. Subsequently, applying the calculated spectrum requirements to the available maritime frequency band, it was validated and discussed in section 4. Finally, a summary of the work and the conclusion drew in section 5.

2. Materials and Methods

2.1. Communication Interfaces for CNPC

The possible communication interfaces for the MASS [7] are shown in Figure 1. The communication interfaces is based on analyses provided through ITU-R WP5B document 5B/841 [2]. The MASS Control Station (MCS) is located at either shore-side

*Corresponding Author: Ho Namgung, Email: ngh2009@mmu.ac.kr

or may be mother ship which controls the MASS for some case. The MCS transmits signals for control and commands, while the MASS sends the MCS her state and also information acquired from onboard sensors - so-called telemetry information. The Vessel Traffic Services (VTS) also communicates with both the MCS and the MASS. As shown in the Figure 1, the navigation by ship can start from voyage planning [8]. At the stage of execution and monitoring, the ship's movement is done in the process of situation awareness through detection and/or identification and sequentially collision avoidance as the result of decision-making by mariner. From viewpoint of safe navigation, the Figure 1 represents the communication interfaces for the safe operations of the MASS. Regardless of the conceptual configuration for the communication interfaces, it is not easy to identify all the information, which is required for a safe navigation. International regulations, The International Regulations for Preventing Collisions at Sea (CORLEGs) [9], The International Convention on Standards of Training, Certification and Watchkeeping for Seafarers (STCW) [10], and The International Convention for the Safety of Life at Sea (SOLAS) [11], can be referred to extract the information required for safe navigation by mariner. However, such regulations describe qualitatively the rules for the collision avoidance among ships and the prevention of any maritime accidents. As presented in Table 1, the international regulations illustrate the articles to be referred to identify the information for a safe navigation. Experiences from mariners can represent the information in detail.

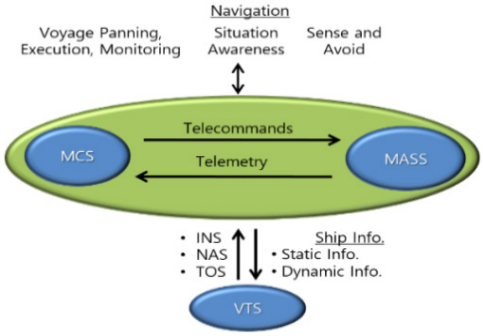


Figure 1: Communication interfaces for safe operations of MASS [7]

Table 1: Regulation related to information for safe navigation

Regulations	CORLEGs	STCW	SOLAS
Articles	Part B. Steering and sailing rules: Rules 5~19	Code A	Chapter V. Safety of Navigation

2.2. Bandwidth for CNPC

To determine the bandwidth for the MASS at sea, a proper frequency band should be identified among available maritime frequency bands. In particular, the CNPC link requires internationally agreed frequency band. This preliminary study provides useful information to find out a proper frequency for the MASS. As described in an Industry Code of Practice, the MASS operations will largely depend on communications systems for control and monitoring of the vessel [12]. It is natural that the CNPC link requires stable and seamless communication. The level of the MASS control will influence on the requirements for the CNPC link. In considerations of the importance of the CNPC link, the two frequency bands, i.e., the Very High frequency (VHF) band, the Ultra High frequency (UHF) band, and the satellite band were being discussed. The present study thus considered the Long Term Evolution - Maritime (LTE-M) for calculating the spectrum requirements of the CNPC.

- LTE-M

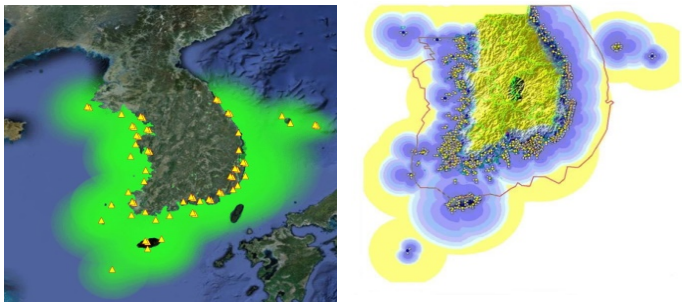
In South Korea, in order to ensure maritime information sharing and communication via access to internet at sea, SMART-navigation project is conducting [13],[14], and [15]. It aims to establish high speed wireless maritime network that covers 100km from offshore through ocean environment conscious cell planning. To facilitate a seamless data transfer various services, therefore, LTE-M is constructing. It is 700 MHz of the UHF band, around the harbor for a variety of services, depending on the strategy to a minimum 1 Mbps transfer rate communication network with the appropriate volume up to 100 km. It enables the free exchange of multimedia data between ships and ship and shore and free use of internet in port and coastal area. In addition, the frequency band of The Downlink (DL) and the Uplink (UL) have10 MHz, respectively.

2.3. AIS

Automatic Identification System (AIS) [16] is an automatic tracking system that uses transponders on ships and is used by VTS. It can transmit information between ships and between ships and land station within range of 50 nautical miles (NM) via VHF in the 160 MHz band. Ships fitted with AIS transceivers can be tracked by AIS base stations located along coast line. Information contains the ship's position, course, speed, name, type, destination, and dimension. SOLAS requires AIS to be fitted aboard international voyaging ships with 300 or more gross tonnage (GT), and all passenger ships regardless of size. A total of 42 AIS base station were constructed on the coasts of South Korea. In present study, in order to take into consideration analysis of the ship traffic flow, data was obtained via AIS during period from January 1 to February 24, 2016.

2.4. Coverage cell

As mentioned in section 2.2, LTE-M enables ships to communicate and exchange data within 100 km (i.e., about 50 nm) of coastal waters from the shore. However, LTE-M with high speed rates for maritime communication network is presently constructing. Therefore, density of the MASS per coverage cell was estimated based on the AIS base stations with about 50 nm of detection distance by considering communication distance and domains as shown in Figure 2.



(a) AIS coverage (b) LTE-M coverage

Figure 2: Comparison of AIS coverage and LTE-M coverage [13],[14], and [15]

2.5. Methodology for Spectrum requirements of CNPC

It was utilized based on '*Characteristics of unmanned aircraft system and spectrum requirements to support their safe operation in non-segregated airspace*' [17] due to non-existence of methodology for spectrum requirements of the MASS's CNPC. In order to calculate spectrum requirements of the CPNC, four step has been composed as presented in Figure 3: (i) analyzed the information exchange rates per the MASS based on the CNPC; (ii) estimated the density of the MASS in coverage cell (e.g., AIS base stations area) per hour; (iii) calculated the spectrum requirements for each type of coverage cell based on (ii) and the frequency efficiency by LTE-M; and (iv) multiplied the reuse factor for each cell type, then determined a total of the spectrum requirements for the CNPC.

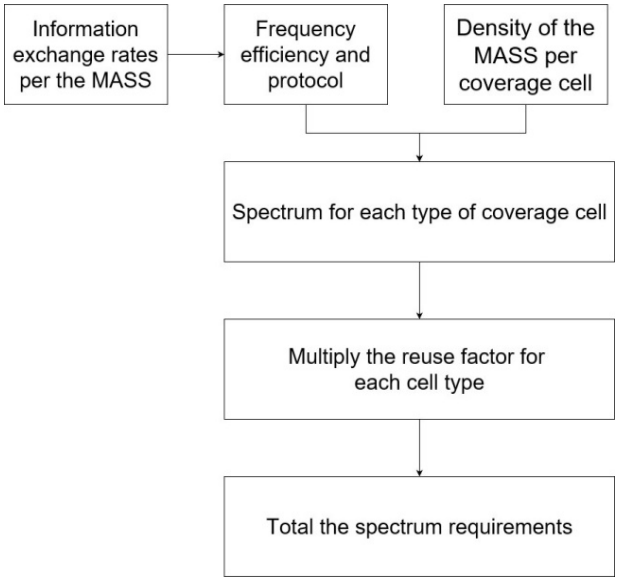


Figure 3: Methodology for spectrum requirements of the CNPC

Calculation methodology of the spectrum requirements for the CNPC of the MASS is shown in equation (1).

$$W_{mass} = K \times B \times M \times R / U \times E \quad (1)$$

Where W_{mass} denotes spectrum requirements, K is reuse factor, B is data rate per the MASS, M is the number of the MASS per cell, R is redundancy factor, U is utilization factor, E is frequency efficiency.

3. Calculation factor for spectrum requirements of CNPC

3.1. Data rate for CNPC (B)

Based on the communication interfaces, a basic data rates is provided shown in Table 2. To calculate the basic data rates, the information required to ensure a safe navigation has been extracted by taking the international regulations into account as mentioned in section 2.1. The DL data is transmitted from MCS or VTS to the MASS, while the UL data is delivered to the MCS or the VTS from the MASS. From Table 2, each data channel needs at least 769.4 Kbps for the UL, and 78.3 Kbps for the DL.

Table 2: Basic data rates for a safe navigation

Telecommand & Telemetry		VTS Communications				Sensing & Collision Avoidance				
		Voice Relay		Data		AIS		RADAR	Target Tracking	Image, Video
DL	UL	DL	UL	DL	UL	DL	UL	UL	UL	UL
0.1742	0.262	64	64	4.53	0.0789	9.6	4.8	100	0.296	600
Total DL(Kbps)		78.3								
Total UL(Kbps)		769.4								

3.2. Density of the MASS per coverage cell (M)

According to the development trend all over the world, the MASS aimed to launch short autonomous runs in 2020 and completely unmanned ships on the high seas by 2025. However, no one knows how many of the MASS would be navigated at sea. Therefore, this study supposes that the MASS would be operated on behalf of present navigating ships at sea. At this time, considering the traffic flow of navigating ships in sea area of South Korea, the number of the MASS was determined.

Among sea areas of South Korea, the Janganseo sea area including the Incheon sea area (the Incheon port, the Pyeongtaek port, and the Daesan port) was considered for analyzing the density of ship traffic. In particular, Janganseo sea area has a lots of ship traffic due to pilot station, anchorage, and junction of inbound and outbound ships. Figure 4 presents the detection range of the Incheon sea area.

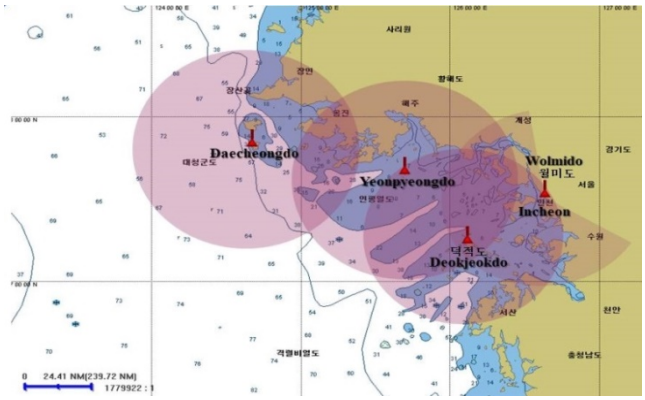


Figure 4: AIS coverage in the Incheon sea area

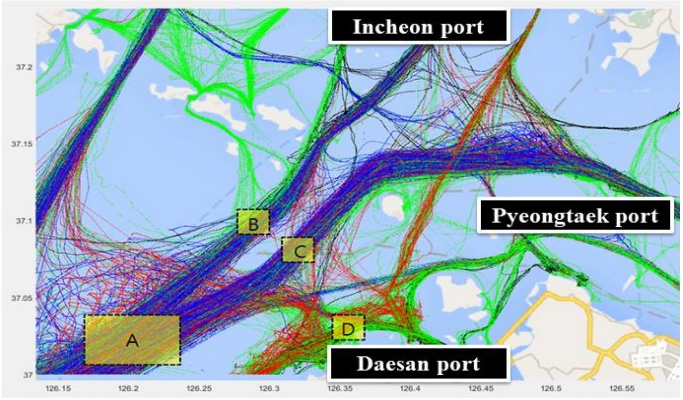
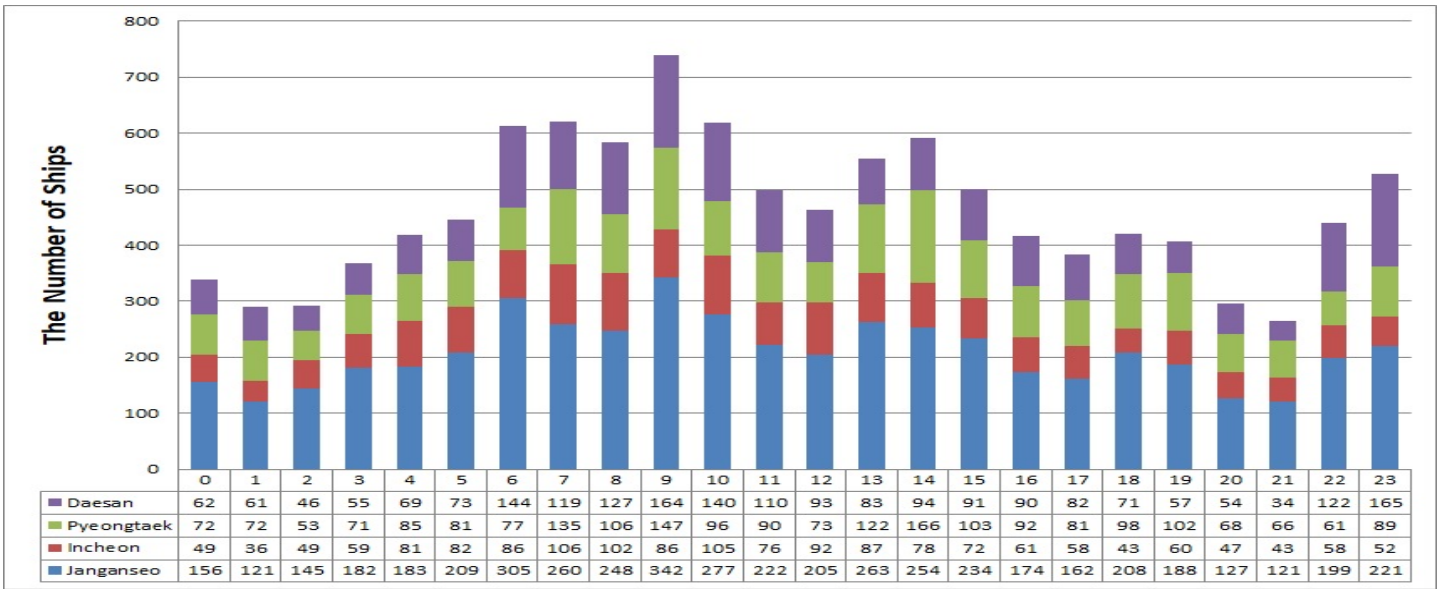


Figure 5: Distribution of ship traffic in the Incheon sea area



	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Daesan	62	61	46	55	69	73	144	119	127	164	140	110	93	83	94	91	90	82	71	57	54	34	122	165
Pyeongtaek	72	72	53	71	85	81	77	135	106	147	96	90	73	122	166	103	92	81	98	102	68	66	61	89
Incheon	49	36	49	59	81	82	86	106	102	86	105	76	92	87	78	72	61	58	43	60	47	43	58	52
Janganseo	156	121	145	182	183	209	305	260	248	342	277	222	205	263	254	234	174	162	208	188	127	121	199	221

Figure 6: Distribution of ship traffic per hour in the Incheon sea area

The number of ships was identified by analyzing the ship traffic statistics by time of day based on traffic density. At this time, it was measured the frequency of ship traffic during the most frequent hours to determine the number of ships per cell. As shown in Figure 5, the Janganseo sea area is distributed as follows: (a) Janganseo sea area; (b) Entrance sea area of Incheon port; (c) Entrance sea area of Pyeongtaek port; and (d) Entrance sea area of Daesan port. Figure 6 shows distribution of ship traffic per hour in the Incheon sea area.

During period from January 1 to February 24, 2016, the highest frequency of ship traffic was at 09:00 hours. The number of ships in the Janganseo sea area (a), the number of entrance sea area of Incheon port (b), the number of entrance sea area of Pyeongtaek port (c), and the number of entrance sea area of Daesan port (d) were 342, 86, 147, and 164, respectively. The Sum of (a), (b), (c), and (d) was 739. And these values were divided by observation period (i.e., from January 1 to February 24) for obtaining the average number of ships in the Incheon sea area. As a result of calculation, 13.4 were measured per hour.

3.3. Redundancy (R) and utilization factor (U)

The Redundancy Factor is information that adds additional information to the original information to ensure accuracy during data transmission. Utilization Factor is the ratio used for each function. Table 3 presents the values of the redundancy factor and utilization factor according to [17].

Table 3: Value of the redundancy and utilization factor

Factor	Telecommand & Telemetry	VTS Communication	Sensing & Collision Avoidance		
			Target tracking	RADAR, AIS	Video, Image
R	2	2	2	1	1
U	0.5	1	0.5	1	0.5

3.4. Frequency efficiency (E)

As the mobile communication network evolves, the frequency efficiency improves. The frequency efficiency refers to the amount of data bits that can be processed for 1 second at a frequency of 1 Hz in a cell. As the frequency efficiency increases, the frequency capacity increases. In order to calculate the required mobile broadband the spectrum requirement within the next 10 years in South Korea, we analyzed the devices that can use mobile communication network within the present or the next 10 years, and estimated frequency efficiency by year [18]. Table 4 shows the annual frequency efficiency according to the LTE technology. According to [18], the present study measured the spectrum requirements by using the frequency efficiency 1.30 in 2020.

Table 4: Prospect of frequency efficiency for LTE

	Prospection				
LTE	2016	2017	2018	2019	2020
	0.97	0.97	1.08	1.19	1.30

3.5. Frequency reuse factor (K)

It is a technology used for the allocated frequency to increase the frequency utilization efficiency. In the case of mobile communication, the service area is divided into a plurality of cells, adjacent cells use different frequencies, and non-adjacent cells use the same frequency to increase frequency utilization efficiency. Table 5 shows the frequency reuse factor according to the shape of the cell.

Table 5: Frequency reuse factor according to the shape of the cell

Cell type	Cell radius (Km)	K applied
A	65	7
B	157	4
C	315	3
D	480	3

We thus applied the cell radius 65 km of cell type ‘A’ due to taking into consideration AIS coverage and LTE-M coverage as presented in section 2.4.

Table 6: Spectrum requirements for CNPC of the MASS

Factor	Telecommand & Telemetry		VTS communications				Sensing & Collision Avoidance				
			Voice Relay		Data		AIS		RADAR	Target Tracking	Image, Video
	DL	UL	DL	UL	DL	UL	DL	UL	UL	UL	UL
K	7	7	7	7	7	7	7	7	7	7	7
B (bps)	174.2	262	64,000	64,000	4,530	78.9	9,600	4,800	100,000	296	600,000
M	13.4	13.4	13.4	13.4	13.4	13.4	13.4	13.4	13.4	13.4	13.4
R	2	2	2	2	2	2	2	2	1	2	1
U	0.5	0.5	1	1	1	1	1	1	1	0.5	0.5
E	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
Hz	50,276.8	75,617.2	9,235,692.3	9,235,692.3	653,713.8	11,385.8	1,385,353.8	692,676.9	7,215,384.6	85,430.1	86,584,615.3
Total DL			11.3 MHz								
Total UL			103.9 MHz								

4. Results and Discussion

4.1. Results

Spectrum requirements of the CNPC was measured according to equation (1). As shown in Table 6, total DL, which is transmitted from MCS or VTS to the MASS, was about 11.3 MHz, and total UL, which is delivered to MCS or the VTS from the MASS was about 103.9 MHz. Accordingly, a total of the spectrum requirements of the CNPC was about 115.2 MHz. At this time, the number of the MASS navigated in further was considered as 13.4 according to section 3.2. Therefore, total DL and total UL would be spent on about 0.8 MHz and about 7.8 MHz, respectively, in case of using a single MASS with the transmit rate 0.8 Mbps.

In terms of the three category: (i) Telecommand and Telemetry; (ii) VTS communications; and (iii) Sensing & Collision avoidance, the amount of the utilization rate out of the total spectrum requirements was observed as presented in Tables 7 and 8. In the utilization rate of the DL, it was indicated that the utilization rate of the VTS communications was higher than other category. On the other hand, in case of the UL, Image and Video were observed as the highest utilization rate. In particular, the total of DL and UL measured from Telecommand and Telemetry was the smallest when comparing of other category.

Table 7: Utilization rate of DL

Telecommand & Telemetry		DL	0.44 %
VTS communications	Voice Relay	DL	81.55 %
	Data	DL	5.77 %
Sensing & Collision Avoidance	AIS	DL	12.24 %
Total DL			100 %

Table 8: Utilization rate of UL

Telecommand & Telemetry		UL	0.07 %
VTS communications	Voice Relay	UL	8.88 %
	Data	UL	0.03 %
Sensing & Collision Avoidance	AIS	UL	0.68 %
	RADAR	UL	6.96 %
	Target Tracking	UL	0.08 %
	Image, Video	UL	83.3 %
Total UL			100 %

4.2. Discussion

In section 4.1, it could be seen that the spectrum requirements of the CNPC and the utilization rate of the UL and the DL were able to be calculated according to [17]. In this section, securing communication channels for a safe operation of the MASS in terrestrial communication system (i.e., VHF band, UHF band) and satellite communication system via the result of the spectrum requirements of the CNPC were discussed, as well as some general reviews of the calculation method.

VHF band (i.e., marine frequency 156 to 174 MHz) is most suitable for short distance communication with the MASS or mutual communication between the MASS and manned ship, but transmission speed is within about 50 Kbps per channel. In case of applying to the calculated spectrum requirements to the VHF band, obtained transmit rate 0.8 Mbps may be impossible to operate in VHF band, and bandwidth, for operating the number of the MASS presented in section 4.1, was not sufficient.

As described in section 2.2, the LTE-M used 700 MHz of the UHF band has transfer rate as a minimum 1 Mbps in bandwidth of the UL and DL, 10MHz , respectively. Applying the calculated spectrum requirements of the CNPC to the LTE-M, the number of the MASS, as mentioned in section 4.1, would not be operated at the same time due to excess of the bandwidth. Whereas, a single MASS with the transmit rate 0.8 Mbps can be operated within bandwidth of the LTE-M due to less than 10 MHz of the DL and the UL, respectively.

According to MUNIN project carried out from 2012 to 2015, they noted that the ship control systems should be designed to operate with different levels of communication quality. Three levels are suggested with respectively 4 Mbps, 128 Kbps and low bandwidth rendezvous services via state of the art satellite communication [4],[5]. In addition, Global Xpress (5 to 50 Mbit/s) was considered for available satellite communication link of the MASS. State of the art satellite communication is possible to obtain required bandwidth according to the user's demand. Therefore, as a result of applying the spectrum requirement of the CNPC to the satellite communications, it may be possible to operate the single MASS or the several MASS. Consequently, the satellite communication link may be more efficient than the terrestrial communication link because a number of the MASS can be operated at sea under control of the MCS due to securing the spectrum requirement of the CNPC. However, satellites normally have a more restricted link budget than terrestrial communication system as power must be generated from limited area solar panels and also because transmission distances are much longer.

A drawback with the calculation of the spectrum requirements of the CNPC is that it only relied on [17] because calculation methods for the MASS has been not existed at present. For this reason, in order to obtain the accurate spectrum requirements of the CNPC, it is necessary to measure suitable factors for the MASS (e.g., reuse factor, redundancy factor, utilization factor, and frequency efficiency).

5. Conclusion

In this paper, appropriate maritime frequency bands were analyzed and verified amongst communication systems via the calculation result of the spectrum requirements of the CNPC. In order to obtain the spectrum requirements, following steps was carried out: (i) obtained the information transmit rates of the CNPC per the MASS ; (ii) estimated the density of the MASS per AIS cell in the Incheon sea area; (iii) calculated the spectrum requirements using factors (e.g., reuse factor, redundancy factor, utilization factor, and frequency efficiency); and (iv) determined a total of the spectrum requirements of the CNPC. The results of applying to the existing and state of the art communication system represented that satellite communication system is suitable for safe operation of the MASS, in terms of bandwidth, transmit rate, and transmit distance. However, in case of using the satellite communication system, it is imperative to invest huge budget for operation of the MASS.

A further study is required to obtain the appropriate values of the factors based on characteristics of the MASS operation system in order to measure the accurate spectrum requirements of the CNPC, and it is necessary to demonstrate the calculation result in comparison with experiment via communication systems (e.g., VHF band , UHF band, and state of the art satellite) at real sea area.

Conflict of Interest

The author declares no conflict of interest.

References

[1] IMO MSC., "Regulatory scoping exercise for the use of Maritime Autonomous Surface Ships (MASS)," MSC 99th session, MSC.99/INF.22, London, 2018.

[2] ITU-R SG, "Characteristics of unmanned maritime systems," The Meeting of working party 5B, Document 5B/841, Bucharest, Romania, 2015.

[3] IMO MSC., "The IMO regulatory framework and its application to Marine Autonomous Systems," MSC 95th session, MSC.95/INF.20, London, 2015.

[4] Ørnulf Jan Rødseth, Beate Kvamstad, and Thomas Porathe, "Communication Architeture for an Unmmaned Merchant Ship," IEEE OCEANS 2013, Bergen, Norway, June, 2013.

[5] Ørnulf Jan Rødseth, Åsmund Tjora, and Pål Baltzersen, MUNIN Deliverable D4.5: Architecture specification, available at http://www.unmanned-ship.org, MRTK, MUNIN-FP7 GA-No.314286, Aug. 2013.

[6] Marko Höyhtyä, Jyrki Huusko, Markku Kiviranta, Kenneth Solberg and Juha Rokka, "Connectivity for Autonomous Ships: Architecture, Use Caes, and Research Challenges," 8th International Conference on ICT Convergence, Oct, 2017.

[7] Ho Namgung, "A Study on Concept Design of Communication System for Safe Navigation of Unmmaned Surface Vehicle," Master dissertation. Mokpo National Maritime University. Mokpo, 2016.

[8] Nathaniel Bowditch, *The American Practical Navigator*, 2002 Bicentennial Ed. Pub. No.9, National Imagery and Mapping Agency, Bethesda, Maryland, 799, 2002.

[9] IMO, *The Convention on the International Regulations for Preventing Collisions at Sea, 1972,* 2010.

[10] IMO, *The International Convention on Standards of Training Certification and Watchkeeping for Seafarers, 1978*, *as amended*, Consolidated Ed. 2011.

[11] IMO, *The International Convention for the Safety of Life at Sea*, Consolidated Ed. 2012.

[12] Society of Maritime Industries UK, *An Industry Code of Practice*, A Voluntary Code Version 1.0, 2017.

[13] Ministry of Oceans and Fisheries, *Report of planning Research for the Development of Korea e-navigation*, 2013.

[14] Ministry of Oceans and Fisheries, *National Strategy for the Development and Implementation of e-navigation*, 2013.

[15] Korea Institute of S&T Evaluation and Planning, *The Report of the Preliminary Feasibility Study on the SMART-navigation project*, 2015.

[16] ITU-R, "Technical characteristics for an automatic identification system using time division multiple access in the VHF maritime mobile frequency band," Rec. ITU-R M.1371-5, 2014.

[17] ITU-R, "Characteristics of unmanned aircraft systems and spectrum requirements to support their safe operation in non-segregated airspace" Rep.2171-5, FEB. 2009.

[18] Korea Communication Commision, "A study on additional spectrum forecast for mobile broadband," 2011.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 297-305 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

BPMN4 Collaboration: An Extension for collaborative Business Process

Leila Amdah*, Adil Anwar

Mohammed V University in rabat, Siweb Team Researsh, EMI, Morocco

ARTICLE INFO

Article history:
Received: 15 September, 2019
Accepted: 07 November, 2019
Online: 12 December, 2019

Keywords:
Business Process modeling
Collaborative business process
BPMN2 Modeler
BPMN Extension Mechanism
BPMN4Collaboration Meta-model

ABSTRACT

Nowadays the collaboration of organizations, systems or services becomes more and more answered in practice. Indeed, organizations try to expand their service, other to satisfy the needs of their customers. Thus, the need to work collaboratively with other partners whether internally or externally becomes trivial. However, process modeling in a collaborative environment becomes heavy because of the diversity of information exchanged and the complexity of shared task. Certainly, the BPMN is the most answered standard for the modeling of business processes, moreover it offers a collaborative diagram that allows to model collaborative processes, but later we will see that its capabilities are very limited. For that, we try in this work to create a collaborative business process modeling language that allows to better model these types of processes and respects some aspects of a collaborative environment. We used the BPMN extension mechanism to create our own language that we named «BPMN4Collaboration».

1. Introduction

The ability to collaborate has always been of vital importance to businesses to businesses. Combining the resources and talents of a group of people together is the best way to solve complex problems and tasks. Collaborative work is the work done jointly by several people who pool their knowledge and skills, organize and coordinate their actions to achieve a result for which they are collectively responsible. Of course, the collective work is not decreed overnight in a company. It is the result of a common desire gradually implemented around a common project. The success factors of this approach are: the benevolence of the various stakeholders, the clarity in the collaboration and the application of the principle give and take. Collaborative work is one of the major challenges of project management. Until now, designers use many languages to model collaborative business process (CBP), such us: XPDL, BPMN, WS-CDL, BPDM, WSFL, XLANG, BPSS, ebXML, WPDL etc.

Business processes are by nature collaborative processes. The modeling of collaborative systems is at the center of current management's concerns and therefore constitutes the main issue of modeling efforts.

Collaboration is viewed as set of exchange between a collection of participants and their processes. We can say that there are two types of collaboration: (1) Integrated Collaboration, which means that business and application logic are integrated into a single monolithic system used by all participants; (2) External Collaboration that takes place between two or more autonomous systems. Practically, collaborative process modeling are very used [1, 2] and software engineering results have shown that collaboration can increase quality and efficiency significantly.

The main goal of a collaborative process is to define the choreography of the interactions between the partners involved in this collaboration. So it's a process that occurs between two or more collaborators and consists of two parts: a public interface and several private interfaces corresponding to each partner figure 1.

- Private process: it's composed of confidential information and is developed internally.
- Public process: it is a partial image of a private process. It consists of all the activities that can be viewed by the other partners.
- Collaborative Process: consists of activities shared among participants during a collaboration.

During a collaboration collaborative process can represent inter-organizational process (between two internal services of the organization) or extra-organizational process (with an external partner). There are several independent identities that represent several decision makers which means autonomy of collaborators; several collaborative activities which means sharing resources and

*Amdah Leila, Morocco, leila.amdah@gmail.com

communication between actors; sharing of responsibility and decisions; finally, a confidentiality and security of resources and data. We can say that all these requirements are essential for a successful collaboration.

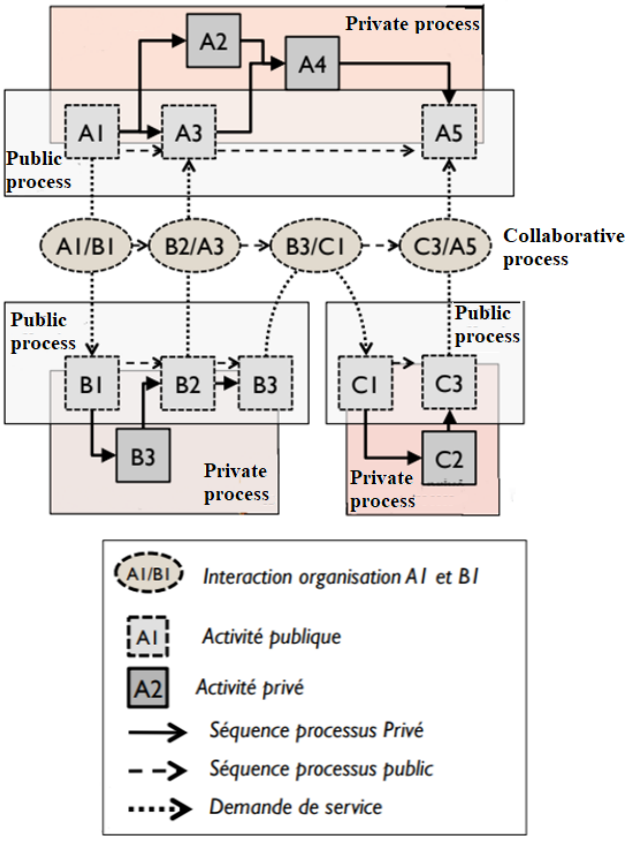


Figure 1: Collaborative Process: Private, public and collaborative process [3]

In fact, BPMN is a business process modeling method that is based on graphical notation and can be used to model collaborative processes. So, we directly think about using BPMN for modeling our collaborative processes. Nevertheless, it still represents some restrictions on the description of the data exchanged between the participants and specification of the roles of the collaborators.

To address the lack of a clear and comprehensive vision of CBP. First we try to define a set of collaborative concepts that are primordial to in collaborative environment namely: confidentiality, traceability, communication, decision making and task monitoring, as present in section 4. Second, we developed a collaborative modeling language with a rich graphic notation corresponding to these concepts of collaboration. Our language named BPMN4Collaboration is an extension of BPMN which is clearly defined because it respect the graphical syntax of the BPMN and finally, offer new concepts for collaborative environment. Our language not only add new notation useful for CBP but also help in the evolution and updating of these processes.

The remainder of this paper is structured as follows: section 2 present background on business process model and notation (BPMN) also the shortcoming of this standard. Section 3 is a related work on BPMN extensions. Section 4 presents an overview on collaborative business process and the central concepts that describe this environment. Section 5 describe the proposed approach which is extending BPMN 2.0 with new collaborative concepts. Section 6 is an illustration to our approach to better understand the benefit of the extension. Finally, we concluded the paper in section 7.

2. Business process modeling

Business Process Modeling is a method of representing the activity of an organization to respond more quickly to customer demand and corporate goals. Thus, it is an organizational management that systematizes the daily tasks of the organization and allows a process to be better understood by the different partners, be it manager, process owner or IT engineering who will implement the system. Business process modeling languages are a great help to model these processes, they are designed for several purposes so it is very difficult to categorize them into one dimension. However, we can distinguish four major scientific traditions of language [4]: the traditional process modeling languages that are typically non-formal, the workflow modeling languages that describe a workflow, they are formal and executable, the integration languages processes and object-oriented languages.

2.1. Business Process Model and Notation (BPMN)

Indeed, BPMN (Business Process Modeling Notation) as Business Process Modeling Language [5] is an important OMG standard that meets the requirements of standardization. Among other popular standards [6, 7, 8], only the BPMN has been widely accepted by the professional world. Thus BPMN is qualified a leader in the field of business process modeling.

The first version of BPMN1.0 was released in 2004 with the purpose of providing a graphical notation to business processes. The last version of BPMN2.0 was published in 2011, this version has brought more elements to the standard and new diagrams (Choreography and collaboration diagram). BPMN 2.0.2, released in December 2013, included only minor modifications in terms of typo corrections. The main advantages of BPMN is:

- Provide a graphical notation understandable by all types of users. Whether it is the business analysts who define the processes or the developers responsible for the implementation, or the employers responsible for the deployment of these processes.
- The possibility of giving life to these processes thanks to BPEL [9] (also known as WS-BPEL) which allows their execution
- Provide an extension mechanism [10] that allows to integrate domain-specific concepts by ensuring the validity of the basic elements of BPMN.

Based in the previous information, BPMN is a base for the representation of business processes. BPMN graphical elements are divided into five basic categories: Flow Objects, Data, Connection Objects, Swimlanes, and Artefacts. Moreover, the BPMN modeling environment consists of three diagrams:

- Processes (Orchestration): These process are called generally workflow or BPMN process, they can be private or public.
- Choreographies: Represent a procedural contract between interacting Participants. Thus, choreography

exists between Pools or participants. It's similar to a process because it consists of activities, Gateways and Events (Figure 2). However, it's different in that the Activities are interactions that represent a set of Message exchanges.

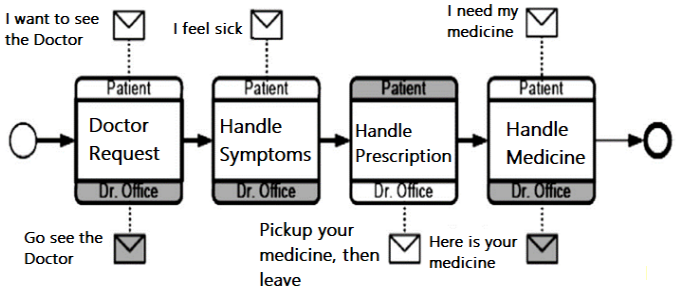


Figure 2: Example of choreographie [11]

- Collaborations: Can include Processes and/or Choreographies: it models the existing interactions between two or more organizations. It contains the pools that represent each participant as well as the message flows. These latter connect the pools (Figure 3). Choreography elements may be shown "in between" the Pools as they bisect the Message Flows between the Pools.

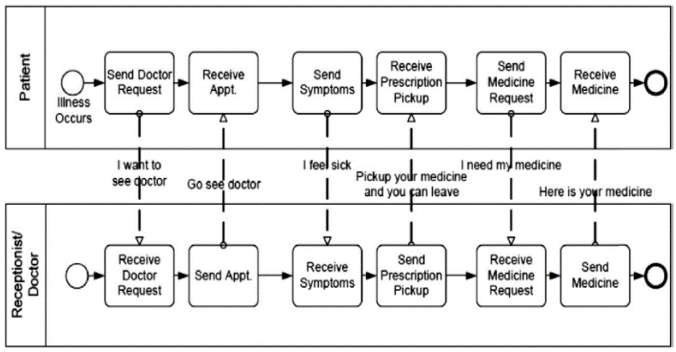


Figure 3: Example of collaboration [11]

2.2. BPMN shortcoming

- Insufficient support for the process state concept. In fact the BPMN does not understand the concept of task progress knowing that this concept is very important to follow the progress of the tasks by collaborators.
- The concept of confidentiality is not taken into account by the BPMN. Confidentiality of information is essential for successful collaboration. The shared information should not be distributed outside or even inside the organization with non-affected persons.
- In PBMN individual participants are represented by Pools so all the activities into this one belong to the same participant. However, a collaboration consist of several participant how can made the same activity. So we need to know if one participant is involved or if there can be potentially many participants involved.
- Of course BPMN represents the exchange of data between several processes. However, this representation remains very limited compared to the need for collaboration. Indeed, in a collaborative environment, the artefact exchange representation is a key point for the success of the collaboration. We must distinguish between the types of these artifacts their level of importance, their version etc .Data objects are represented simply by a symbol; their internal structure is not defined and their main purpose is not specified. However the exchange of data is a very sensitive point when it comes to collaborative work. The type of data and their purpose must be specified to add value to the collaborative process.

3. Related works

Several works benefit from the mechanism of extension of the BPMN to adapt it to their needs. We will present therefore some work that has developed BPMN extension in different ways: in the field of knowledge dimension and sensitive business process model, [12] developed an extension (named BPMN4KM) that link different types of knowledge, information and data to the BPMN model. Other work [13] focuses on the field of clinical pathways and develops an extension (named BPMN4CP) which allows to integrate the concepts of this domain into BPMN model.

Concerning collaborative environments, [14] is interested in the life cycle of the data objects and the information exchanged between the processes during a collaboration. Thus, it develops an extension to model artifact-centric processes by taking inspiration from the existing graphical notations and best modeling practices. As for [15], he is interested in highly collaborative processes by adding new notations to model tasks that are common in this type of process. In [16] the author treats on complex collaborative processes. He presents the limitations of the BPMN concerning the modeling of complex choreographies. So he develops an extension to remedy this problem by adding new concepts, then he validates his work by using Service Interaction Patterns.

In other field the author introduce BPMN4TOSCA[17]. He expands the BPMN to model portable management plans for a standard called TOSCA. In fact, this standard captures cloud application topologies. So the author, defined four extension namely: topology management task, node management task, script task and finally data object specific to TOSCA. Zor and al. [18] propose a BPMN extension for the manufacturing domain to explicitly handle products and resources. [19] Used the BPMN extensibility mechanism to model specific tailoring relationships such as: deletion, local contribution, and local replacement. In addition, the author tries to ensure the consistency of the BPMN models when using tailoring relationships by creating some specific rule. Some works are interested to include the security aspect into BPMN. For example in [20] the author include access control or intrusion detection through a set of new annotations. To introduce security concepts in process models, [21] extend the BPMN by modeling elements such as access control or separation of duty, binding of duty. It names its extension SecureBPMN dedicated to secure business process.

Extension mechanisms tend to develop new concepts for a specific use from a general modeling language, or to enrich already existing elements with new information. The technical structuring of BPMN is based on the concept of expandable layers around a simple element core that provides this extension mechanism. As

you can see, each approach represents a new BPMN extension for a specific domain. While there are some who have approached the collaborative process, none of them has specified an extension that covers the different aspects of collaborative processes. Inspired by some work, we try to extend BPMN with graphic elements that:

- Represent concept of CBP
- Respect the shape of the standard
- Appear the extension as natural as possible

4. Collaborative business process

4.1. Overview

The concept of collaboration is derived from the Latin *collaborare* meaning "to work together" and can be seen as a process of shared creation. It can be defined as a process in which partners share information, resources and responsibilities to jointly evaluate a program of activities to achieve a common goal. Thus it's a process by which one group of entities enhances the capabilities of the each other. Collaboration involves mutual engagement of participants which implies mutual trust, sharing risks, responsibilities, and rewards. A collaborative process occurs, for instance, when a team of experts jointly develops a new product or improve an internal service to satisfy their customers. Such collaboration concerns the association between different organizations by connecting their business processes and exchanging data, messages or resources between their different departments. Thus the development of such collaborative systems becomes a challenge because it requires the linking of several heterogeneous elements of each organization of course by creating a coherent overall system and keeping the internal workings of each organization intact. In addition, this must be done in accordance with the organization's policies, industry guidelines, applicable legislation, technical capabilities and limitations.

Furthermore, organizations must be able to model their requirements so that: its internal processes describe its private activities; its public processes describe how it can interact with others; and its partner's public process describes what it expects from its partner. These requirements can be diverse, ranging from sequencing activities to sharing resources, defining security objectives, etc. On the basis of these requirements, organizations must then have the means to establish a clear, well-defined collaborative process that respects these requirements and the consistency of private and public behavior of partners.

To this end, there is strong interest from organizations in utilizing collaborative modelling concepts to be able to specify the different requirements applicable to their collaboration. On the basis of the resulting collaborative model, companies must then be able to evolve their internal processes or rectify specific requirements of the collaboration contract without affecting the proper functioning of the overall collaborative process as well as partner's process. Finally, the sketched development allows the modeling of the resources exchanged between the partners as well as the flow of relevant activities. To describe and model collaborative business processes, several companies have come up with their own specification languages, such as Microsoft, IBM, SAP, etc.

4.2. Concept of collaboration unit

Previously we have seen that the BPMN is limited with respect to collaboration. In fact, collaboration is not just about sharing resource and exchanging messages between collaborators but it is a very complex operation that requires a lot of effort from participants. Collaborative business process are difficult and their success depend on some requirements:

- Traceability: Sharing implies sharing responsibility of participation and sharing responsibility of results. Indeed, a successful collaboration depends on a job with traceable tasks that allow the smooth running of the process.
- Confidentiality: Confidentiality of information is essential for successful collaboration. The shared information should not be distributed outside or even inside the organization with non-affected persons.
- Communication and sharing of resources is a great challenge, whether it is resources exchanged by collaborators or resources acquired during a meeting in order to fulfill the task.
- Decision-making: any decision made must be shared with collaborators
- Task monitoring and the progress of work: these two points are of a great importance in collaboration work. When sharing resources, it is necessary to assess their stat of progress.

To develop rich collaborative business process modeling. We try to exploit as much as possible the requirements we have mentioned before thus we propose a conceptualization for specifying collaborative business process in the form of a generic BPMN Meta-model for Collaboration (BPMN4Collaboration). Our work is based on the principle of extending the basic concepts of BPMN namely Activity, DataObject, SequenceFlow etc. The extension of the concepts of collaboration on these artifacts is detailed in Figure 4.

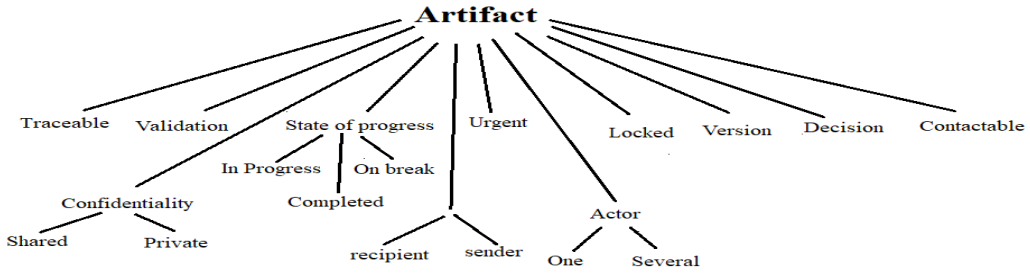
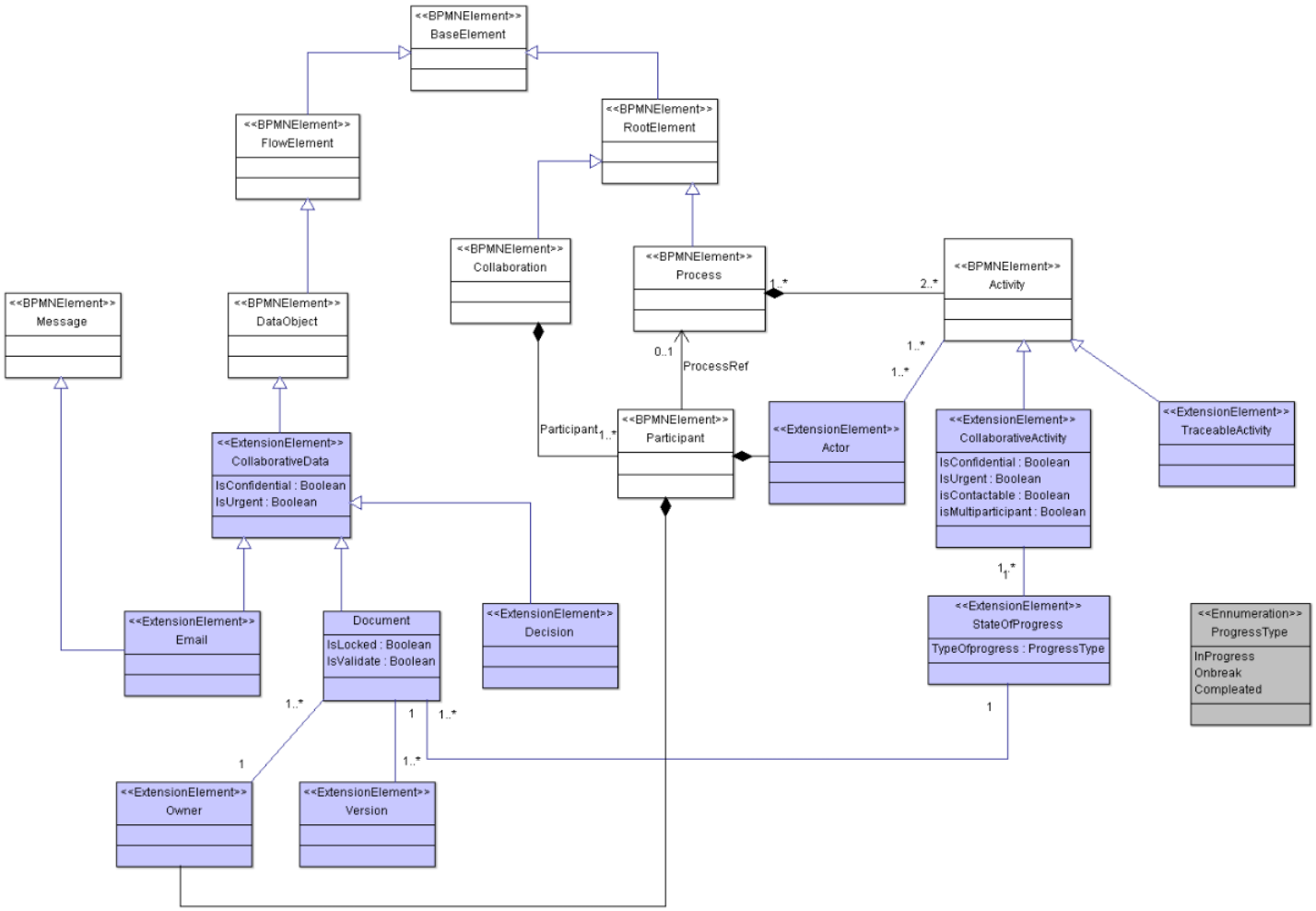


Figure 4: Extension of the concept of collaboration on BPMN artefacts



The concepts discussed in this section will help us later to build a modeling language that better describes business processes in a collaborative environment.

5. BPMN4Collaboration: A BPMN extension including collaboration concepts

5.1. Extended BPMN4Collaboration Meta-model

The metamodel of our extension called BPMN4Collaboration is shown in Figure 5. We try to represent in detail the stereotypes that describe the new concepts of collaboration

CollaborativeActivity: We add the property of privacy, urgency, the mode contactable and multiparticipant by creating this stereotype which is an extension of BPMN Activity. The attribute isConfidential show that Activity can be private or shared. Private activity can be seen only by the participant to which it belongs. Otherwise, shared activity is seen by all participants. The attribute isUrgent specifie that the attribute isContactable specifies that we can contact this activity before the end of the execution of the process. The attribute isMultiParticipantTask specifies if activity is made by several participants. A collaborative activity has also a state of progress.

TraceableActivity: this stereotype add the option of traceability. We can trace the execution status of the activity

Actor: this metaclass specifies that activity has an actor so that it can be on or several actors.

CollaborativeData: extends BPMN Data Object. It has also the attribute isConfidential and isUrgent. A collaborative data can be in the form of: Email, Decision or Document

Decision: this metaclass is a type of collaborative data, it must be an output of an activity. It's always connected with an activity and a Message Notification event. A decision must be notified to all participants before the end of the process.

Document: this metaclass is a type of collaborative data, it has two attributes: isLocked to specify that document cannot be modified or updated. isValidate to specify that the manager or responsible of this document has validated it. A Document has also an owner and version and state of progress.

Owner: a document has always an owner which describes the initial state of the document for example decide if the document is locked or not, if it's confidential, etc.

Version: this metaclass specifies the version of the document. A document has inevitably at least a first version and a final version if it's completed.

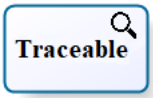
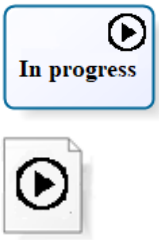
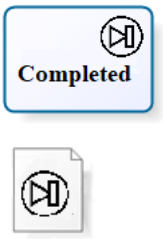
stateOfProgress: this metaclass is linked to the stereotype CollaborativeData and CollaborativeActivity to shown the state of

progress of artifact. It has three states: in progress, on break and completed.

5.2. Graphical Concrete Syntax

We propose an advanced graphical syntax which defines the elements of the new concepts of BPMN4Collaboration in Table 1 below:

Table 1: Extended element of BPMN

Abstract syntax	Graphical concept	Description
Multi participant Activity		Activity made by many participants
Contactable Activity		This activity can be contacted at any time by an external partner to be able to make changes before the end of the process
Traceable Activity		A task is said trackable to an external organization, if that organization is allowed to trace the execution status of the task
Private Artifact	 	This element is used to specify that the artifact is confidential and cannot be shared with a third party
Shared Artifact	 	Artifact shared between partners
Urgent Artifact	 	the artifact is urgent so its treatment must be immediate
State of Progress of an Artifact		The artifact is in progress
		The treatment of the artifact is on break
		The artifact is completed
Validate Data		Data validate by the responsible of the data
Locked Data		The exchanged data is locked, so it cannot be modified or updated.
Decision Data		Set of decision made by the partners before the end of the process that must be sent to the relevant stakeholders.
Data Rules		Data Rules which are mandatory before performing the task
Version of Data		Specifies that Data has a version « X »

5.3. Discussion

Through this modeling, we have been able to achieve an extension that satisfies the collaboration concepts mentioned in section 4. Indeed:

- Traceability is modeled by "traceable activity"
- confidentiality lies in defining « private artefacts », « public artefacts » and « locked data »
- Communication and exchange of resources is a trivial point that is represented by all of these new elements and specially « contactable Activity ».

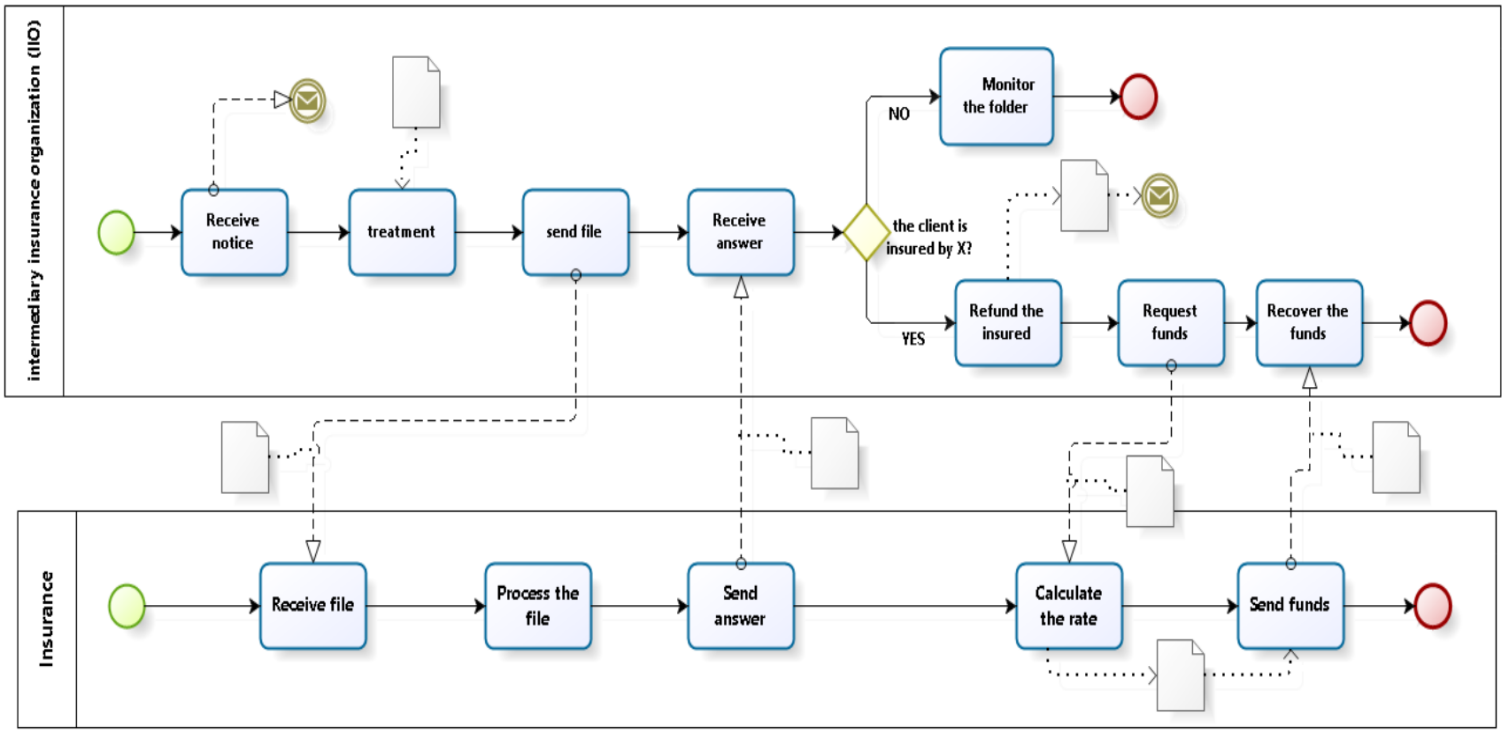


Figure 6 : Example of collaborative BPMN within an insurance

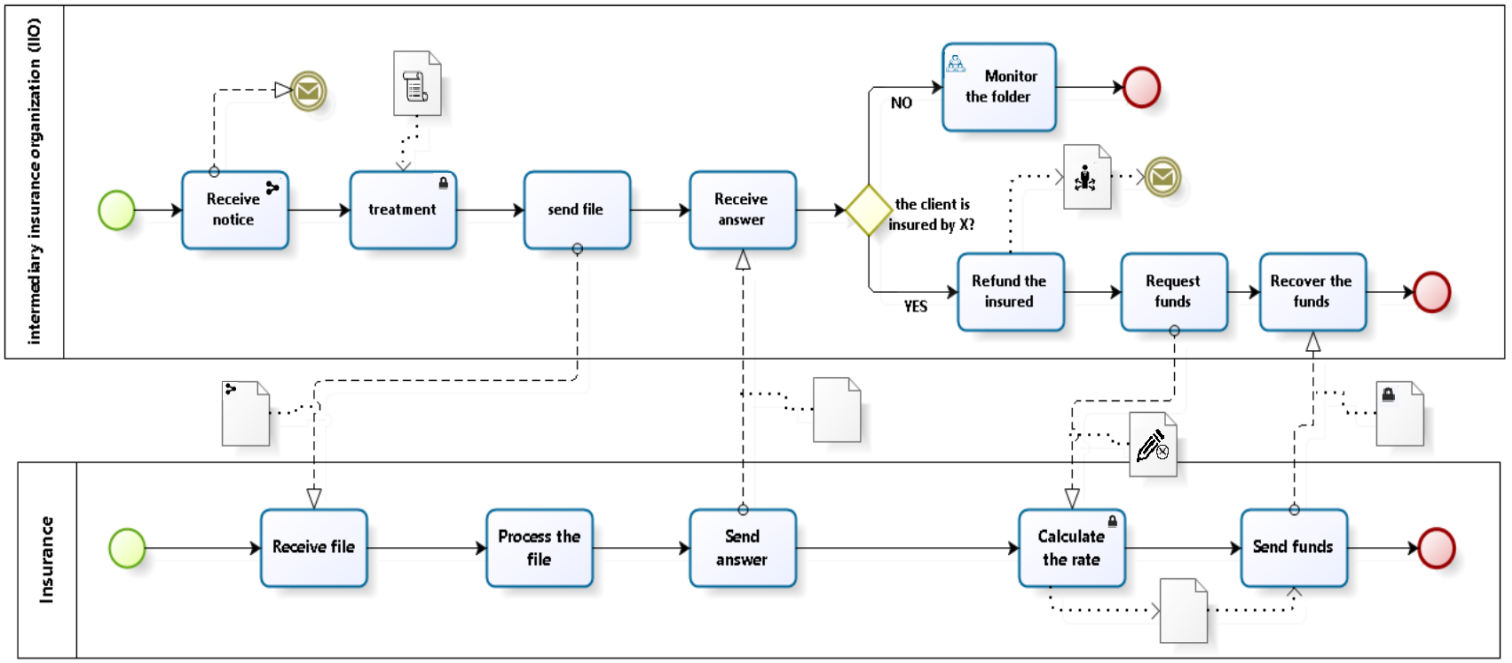


Figure 7: Example of collaborative BPMN conform to collaborative business process language

- Decision-making is ensured by « Decision Data » and « Data Rules ».
- Task monitoring and progress of work are provided by the elements « state of progress artefact », « urgent artefact » and « validate data ».
- Finally, the collaboration is more precise by "multi-participant Activity". The latter allows to emphasize the fact that task can be performed by several participants and not by the actor specified by the model BPMN.

6. Demonstration

6.1. Illustrative example

Our example illustrates a collaboration between several organizations in the insurance field. Indeed, the collaboration is between the IIO and several insurances;

- Intermediate Insurance Organism (IIO): It is an organization that mediates between national and international insurance.
- Insurance: In our example we call our insurance "insurance X" as the IIO collaborates with several insurance at the same time.
- Description of the scenario: A person having an accident in a foreign country, the IIO will receive his notice of damage then will contact his insurance on his country to take charge of the material and corporal damage.

The process begins at the IIO as seen in Figure 6. As soon as the IIO organization receives a notification of damage from his client (the victim). It notifies the concerned insurance (named in this example X) by sending a notification message. Before the process of the file some rules of collaborations must be checked by collaborators (these rules specifies that insurance X is in agreement with the organism). The file is then sent to the corresponding insurance which deals with it in turn and sends an answer to the organism. The process continues, if the client is not covered by insurance X, the IIO supervise the filed, then, the process ends. Otherwise, if the victim is insured by the insurance X the IIO reimburse the client. Of course this decision is sent to all collaborators. The IIO requests the funds from the insurance company, the latter makes its calculations and sends them to the organism, thus, the process of insurance X ends. Finally, the process of the organism ends after having received funds. In this process, there are several actors, namely: the client (or the victim), the intermediary organization (IIO) and the insurances that collaborates with the IIO. The exchange of data between these participants is done either by sending documents by email or post.

We notice that this modeling is not precise since a quantity of information is not specified on the model.

Figure 7 demonstrate the evolved BPMN extension by presenting the new concepts of collaboration: Shared activity like “receive notice” is a public activity shared between client and IIO. Private activity like “calculate rate” is private, it’s done inside the insurance and can’t be seen by other participant. To supervise the progress of the treatment of the file, the IIO must keep in touch with this task. Then, the activity “process the file” done inside the insurance is traceable by the IOO. “Monitor the folder” is an activity made not only by the IIO, but by the insurance and the client, so it’s a multi-participant task.

Concerning the exchange of data, in our extend model, data Objet are more clear. For example we can distinguish the private data (which keep the confidentiality of data). Rules date are input for the activity “treatment”, in fact, this activity can’t be performed if these rules are not satisfied. The integrity of data is clearly defined in the model (editable and locked data). And finely the decision data, when the IIO decide to refund the insured, all the participant (client, insurance) must be informed by a notification message .

6.2. Tool demonstration

For our demonstration we use Eclipse BPMN2 modeler which is built on the Eclipse Plug-in Architecture. This graphical modeling tools provide extension points that help to customize and edit the appearance of editor by using the extension mechanism of BPMN2.0.

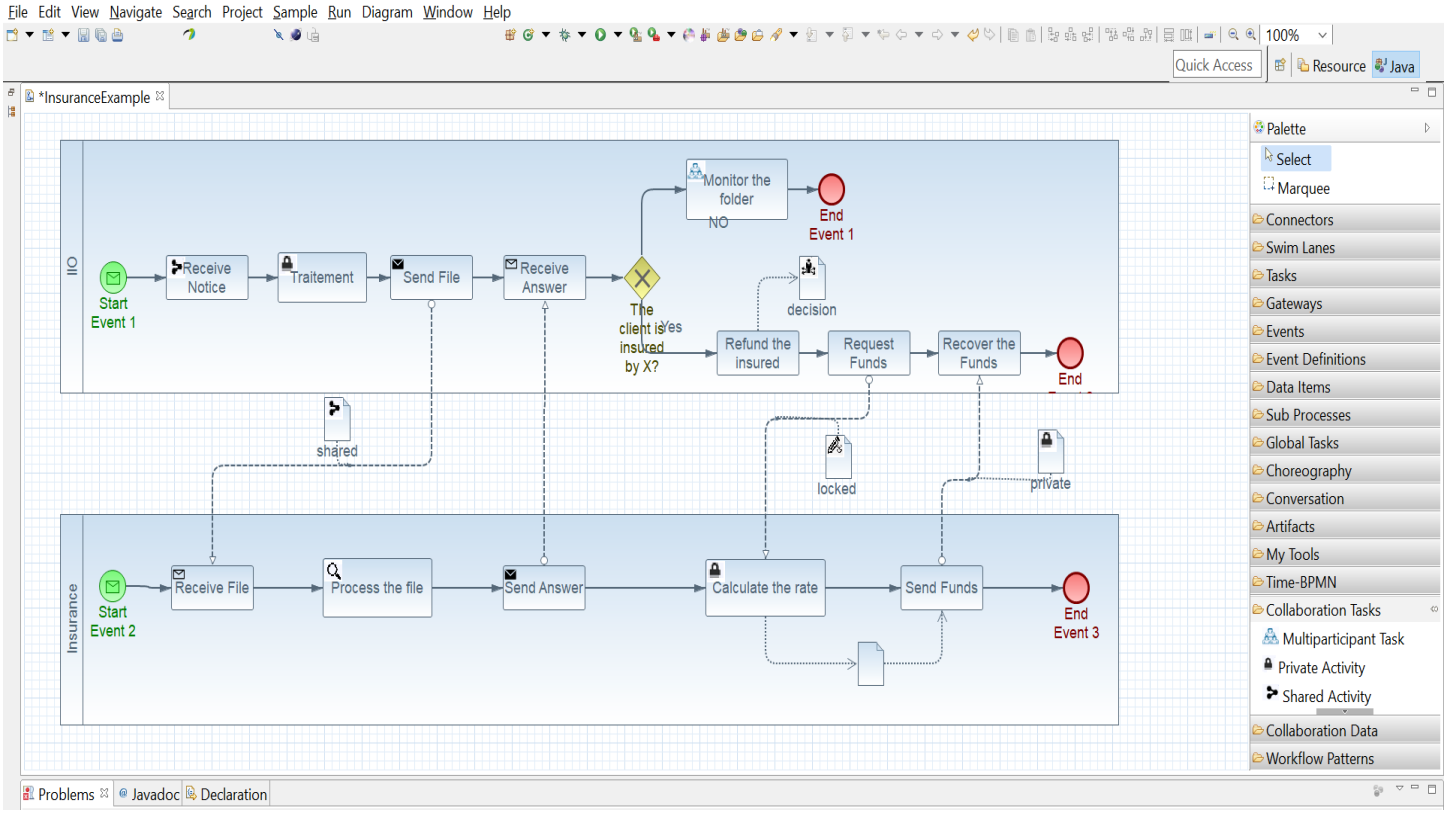


Figure 8: Screenshot from Eclipse that present Extension for Eclipse BPMN2 Palette

We developed a specific plug-in that represent the aspects of collaboration which are treated by our approach. As we have explained is an extension of the Eclipse BPMN2 Modeler plug-in [22]. It allows us to integrate new attributes, properties and new graphic elements adapted to our needs.

As we have extended in our work BPMN Activities and Data Objects. Now the BMPN2 palette include this new graphical element.

An example of a BPMN4Collaboration diagram and our proposed extension is depicted in Figure 8. We can see two tabs added in the palette "Collaboration Task" and "Collaboration Data". They contain the new graphical collaboration syntaxes explained by our approach.

7. Conclusion and future work

The modeling of collaborative processes is handled by several works and languages, namely the BPMN which is the standard most used by collaboration tools and companies. However, even the BPMN recognizes limitations in modeling collaborative processes and especially the modeling of data exchange that are presented in a general and unclear way. We have exploited the capacity of BPMN 2.0 extension mechanism which is specially designed to represent extensions of its standard elements to create an extension for collaborative environments [23]. We tried to add new concepts useful for BPMN modeling and we gave more attention to the representation of the data exchanged between the collaborators. To create this extension we began by collecting some aspects useful in modeling collaborative process, after we create our proper metamodel by extending BPMN metaclasses and adding new stereotypes. Then, we gave a graphical syntax to this new elements, also we have defined a textual semantic for the new extended element. We can say that our extension is a useful for modeling collaborative processes by providing more precision on the shared elements between partners.

Collaborative business processes are very complex and our extension only allows to treat a small part of this complexity. For that we will try in our next work: to enrich our extension even more by adding other aspects like security which is primordial in a collaboration environment. We will define a complete collaborative business process modeling language by adding semantics to our language. As we know, BPMN can be mapped to many executable languages to be implemented, so we will explain exactly how to implement our new elements, and how they can be implemented for example in existing BPMN tools such as JBoss Jbpm [24] or BonitaSoft [25] or others. And finally, the execution of these elements will be treated.

References

[1] Rittgen, P.: Collaborative Modeling – A Design Science Approach. In: Proc. HICSS ’09. (2009) 1–10

[2] J.Mendling, J.C.Recker,., Wolf, J.: Collaboration features in current bpm tools. EMISA Forum 32 (2012) 48–65

[3] Cours Angnès Front «Systèmes d’Information Conception et Processus»

[4] H. Mili, G. B. Jaoude, É. Lefebvre, G. Tremblay, and A. Petrenko, « Business Process Modeling Languages: Sorting Through the Alphabet Soup », p. 54.

[5] Business Process Modeling Notation Specification, OMG Final Adopted Specification, February 2006, dtc/06-02-01 (http://www.bpmn.org/Documents/OMG Final Adopted BPMN 1-0 Spec 06-02-01.pdf).

[6] Mayer, R.J., Menzel, C.P., Painter, M.K., deWitte, P.S., Blinn, T., Perakath, B.: IDEF3 Process Description Capture Method Report, Knowledge Based Systems, Inc., 1997.

[7] Eriksson, H.E., Penker, M. Business Modeling with UML: Business Patterns at Work, Wiley, 2000, ISBN 978-0-471-29551-8.

[8] Scheer, A. W: Architecture of Integrated Information Systems: Principles of Enterprise Modeling. Berlin et al.. (1992).

[9] Web Services Business Process Execution Language Specification. <http://docs.oasis-open.org/wsbpel/2.0/OS/wsbpel-v2.0-OS.pdf.>

[10] Stroppi, L.J.R., Chiotti, O., Villarreal, P.D.: Extending BPMN 2.0: Method and Tool Support. Business Process Model and Notation, 59- 73 (2011)

[11] M. von Rosing, S. White, F. Cummins, and H. de Man, « Business Process Model and Notation—BPMN », in *The Complete Business Process Handbook*, Elsevier, 2015, p. 433-457.

[12] M. Ben Hassen, M. Turki, and F. Gargouri, « A BPMN Extension for Integrating Knowledge Dimension in Sensitive Business Process Models », in Information Systems, vol. 299, M. Themistocleous and V. Morabito, Éd. Cham: Springer International Publishing, 2017, p. 559 578.

[13] R. Braun, H. Schlieter, M. Burwitz, and W. Esswein, « BPMN4CP: Design and implementation of a BPMN extension for clinical pathways », 2014, p. 9 16.

[14] N. Lohmann and M. Nyolt, « Artifact-Centric Modeling Using BPMN », in Service-Oriented Computing - ICSOC 2011 Workshops, vol. 7221, G. Pallis, M. Jmaiel, A. Charfi, S. Graupner, Y. Karabulut, S. Guinea, F. Rosenberg, Q. Z. Sheng, C. Pautasso, et S. Ben Mokhtar, Éd. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, p. 54 65.

[15] Pedro Antunes and Al. « Modeling highly collaborative process » 2013

[16] G. Decker and F. Puhlmann, « Extending BPMN for Modeling Complex Choreographies », in On the Move to Meaningful Internet Systems 2007: CoopIS, DOA, ODBASE, GADA, and IS, vol. 4803, R. Meersman et Z. Tari, Éd. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, p. 24-40.

[17] O.Kopp, T.Binz, U.Breitenbucher, and F.Leymann « BPMN4TOSCA: A Domain-Specific Language to Model Management Plans for Composite Applications »

[18] Zor, S., Leymann, F., Schumm, D.: A Proposal of BPMN Extensions for the Manufacturing Domain. In: Proceedings of the 44th CIRP Conference on Manufacturing Systems. (2011)

[19] Raquel M.Pillat, Toacy C.Oliveira, Paulo S.C.Alencar, Donald D.Cowan “BPMNt: A BPMN extension for specifying software process tailoring » Information and Software Technology Volume 57, January 2015, Pages 95-115

[20] Rodriguez, A., Fern´andez-Medina, E., Piattini, M.: A BPMN Extension for the Modeling of Security Requirements in Business Processes. IEICE Transactions on Information and Systems 90(4) (2007) 745–752

[21] Brucker, A.D., Hang, I., L¨uckemeyer, G., Ruparel, R.: SecureBPMN: Modeling and Enforcing Access Control Requirements in Business Processes. In: ACM Symposium on Access Control Models and Technologies. (2012)

[22] BPMN2 Modeler. http://www.eclipse.org/bpmn2-modeler/

[23] L.Amdah and A.Anwar : BPMN Profile for Collaborative Business Process . The 5th International IEEE CiSt Congress on Information Sciences and Technology (Cist’18) .Marrakech, Morocco, October 21 – 24 ,2018

[24] BonitaSoft Home page.< http://www.bonitasoft.com/>

[25] JBoss jBPM Home page. < http://www.jboss.org/jbpm>

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 302-306 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

The Impact of Manufactured Sand (M-Sand) as Partially and Fully Replacement of Fine Aggregate in Concrete

Altamashuddinkhan Nadimalla[1,*], Siti Aliyyah Masjuki[1], Siti Asmahani Saad[1], Ean Lee Woen[2], Siti Maisarah Ali[1], Naseer Ulla[3]

[1]*Civil Engineering Department, International Islamic University Malaysia, 53100, Malaysia*

[2]*Institute of Sustainable Energy, Universiti Tenaga Nasional, 43000, Malaysia*

[3]*Civil Engineering Department, Bearys Institute of Technology, 574153, India*

ARTICLE INFO

Article history:
Received: 23 November, 2019
Accepted: 19 January, 2019
Online: 27 January, 2020

Keywords:
flexural test
Impact test
compression test

ABSTRACT

This, research work was to study the potential of M-sand as compared to river sand in concrete, here M-sand is replaced by river sand 0%,45%,50%,55% and 100% in the concrete mix, Mix design is designed as per IS Standards. In this research a mix 1:2.32:2.82 (M20) was considered. The test specimen was casted for 7days, 28 days and 90days. The performance of M-sand was determined by several experiments such as slump test, impact strength test, flexural Strength, and compressive strength test. The results attained from each test states that as M-sand increases the slump value decreases. flexural strength, compressive strength and impact test of concrete at 7 days, 28 days and 90 days is greater at 100% and 50% replacement of M sand by river sand.

1. Introduction

This paper is an extension of work originally presented in conference IEEE 5th International Conference on Engineering Technologies & Applied Sciences, 22- 23 Nov 2018, Bangkok Thailand [1]. In the world, fine aggregate, probably natural sand, composes up to 30% of the volume of concrete, it is around 4 billion tons sand is required to meet annual necessity of concrete production [2]. Increasing in excavation of river sand from riverbeds leads to serious threat to environment [3]. Due to the limiting resources of river sand, alternatively M sand can be used. [1]–[10]. The Manufactured sand is produced by crushing the rocks [5] [10].

The impact of M-Sand on the concrete properties:

Higher split tensile strength, higher compressive strength, & higher flexural strength can be attained by the 50% substitution of of fine aggregate by M Sand [5], [6],[10]. It is due to the angular in shape of M sand gives good bonding between cement and aggregate. Gradual increase in strength can be obtained by 50% replacing river sand by M sand. [4]. Gained higher strength at 60% by substituting the river sand by M sand at various proportions such as 0%, 20%, 40%, 60% and 80% [14]. The concrete flexural strength increased up to 2% and 4.3% by 25% and 100% incorporation of river sand by M sand[15]. Compared to the air curing and standard moist curing, the membrane curing will give the good strength in both river sand and M sand. Addition of Super Absorbent Polymer leads to a significant increase of mechanical properties of the concrete [16].

The M sand properties are very much similar to river sand so nowadays, instead of river sand, M sand can be used in place of river sand[11] , by utilizing M sand in alternative of river sand the workability of concrete will decreases[12], the workability can be gained by adding water reducing admixtures [16].

It was observed that if river sand is completely replaced by M sand, M sand has greater resistance to loss in strength as compared to river sand, if specimen is immersed in chemicals. The Combined replacement of M sand and Marble powder upto 25% by river sand increased the concrete strength [17].

Compressive strength increased by approximately 10% for 28 days of curing, when river sand fully substituted by M sand [2]. It has been recorded that the compressive Strength enhance up to 50% incorporation of river sand by M-sand sand and introduction of 2% to 6% waste plastics [18].

The 100% incorporation of natural sand by M sand, it helps in increase in paste volume as compare to river sand, which is useful to produce self-compacting concrete. Increase in paste volume is

*Corresponding Author: Altamashuddinkhan Nadimalla, Email: altamashk1987@gmail.com

due presence of high fines in M sand, which increases the water demand also [19].

Manufactured sand is suitable as alternative for river sand at affordable cost. It will act as cohesive cement mortar. It helps environment to maintaining the economical balance [3] [19].

2. Objective

- To inspect the properties of M-sand
- To examine the impact of M-sand on fresh and hard properties of concrete.
- To investigate the comparison of the, flexural Strength, compressive strength and Impact sstrength achieved by the cubes and beams in river sand and M-sand.

3. Materials and Its Properties

3.1. Cement

In this research Ordinary Portland Cement (53 grade) was used. As per IS: 8112-1989 Cement Properties are determined as tabulated in Table 1.

Table 1: Cement Properties

Properties		Result	IS code Requirement
Specific gravity		3.05	
Standard consistency		31%	
Setting time	Initial(Minutes)	33 min	30 minimum
	Final(Minutes)	380 min	600 maximum
Compressive strength (MPa)	7 days	32	33
	28 days	46	43

3.2. Fine Aggregates

The river sand was collected from the Mangalore local area which is excavated from riverbeds. The sieve analysis of river sand is tabulated in Table 2 and shown in Figure 3.

Table 2: Sieve analysis of river sand

SL No	Sieve Designation	Percentage Passing	Grading Limits for Zone II Sand (IS383)
1	4.75 mm	99	90-100
2	2.36 mm	87.8	75-100
3	1.18 mm	54.6	55-90
4	600 micron	37.6	35.59
5	3000 micron	12.6	8-30
6	150 micron	6.4	0-10
7	75 micron	2	0-10

From the sieve analysis test the river sand and M-sand fineness value are 3 and 3.45 respectively, both falls under zone-2, it shows the M-sand properties is similar to river sand, M-sand is slightly coarser as compared to river sand. The sieve analysis of M sand is shown in Table 3 and Figure 4

Table 3: Sieve analysis of M- Sand

SL No	Sieve Designation	Percentage Passing	Grading Limits for Zone II Sand (IS383)
1	4.75 mm	100	90-100
2	2.36 mm	91	75-100
3	1.18 mm	61.2	55-90
4	600 micron	48.3	35.59
5	300 micron	27.4	8-30
6	150 micron	13.8	0-10
7	75 micron	3.8	0-10

3.3. Coarse Aggregate

The 20mm down size crushed stone is used in this research. Table 4 and Figure 5 shows the Coarse aggregate sieve analysis.

Table 4: Sieve analysis of Coarse Aggregate

SL No	Sieve Designation	Percentage Passing	Gradation requirement as per IS 383-1970 for grade II	Remark
1	40	100	100	As per IS 383 the sample confirms the graded aggregate
2	20	98	95-100	
3	10	46	25-50	
4	4.75	0.8	1-10	

3.4. Mix Design

In this research work a mix design 1:2.32:2.82 (M20) is considered.

Table 5: Mix Design

Mix Design	Cement (kg/m^3)	Aggregate (kg/m^3)		w/c Ratio	Water (l)
		Fine	Coarse		
1:2.32:2.82	358	829.67	1009.96	0.55	197

4. Fresh Properties of Concrete

4.1. Slump test

The workability of concrete was estimated by using slump test. Figure 6 and Table 6 shows that slightly decrease in the slump value as M sand increases.



Figure 1: Slump test

Table 6: Slump test

Percentage of M Sand	Slump Value (mm)
0%	102
45%	100
50%	99
55%	100
100%	99

5. Hard properties of Concrete

5.1. Compressive Strength test

The tests were carried out on 150mm* 150mm*150mm size of cubes. The outcomes is displayed in Figure 8.

Figure 8 shows the compressive strength for results for 7days, 28 days and 90 days curing, it is detected that the compressive strength increased for 100% and 55% of replacement of M Sand by river sand, 100% and 55% incoporation of M sand can be advised to use as fine aggregate.



Figure 2: Compressive strength test

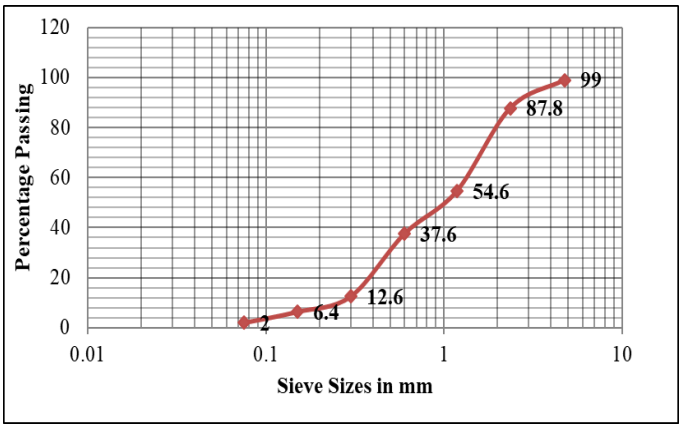


Figure 3: Sieve analysis of River Sand

5.2. Flexural strength test

The tests were carried out on beams of width 150 mm length 150 mm and 70 mm thickness. Figure 9 shows the flexural strength for results 7days, 28 days and 90 days curing, it is observed that the flexural strength increased for 100% and 55% of incorporation of M Sand by river sand, 100% and 55% incorporation of M sand can be advised to use as fine aggregate.

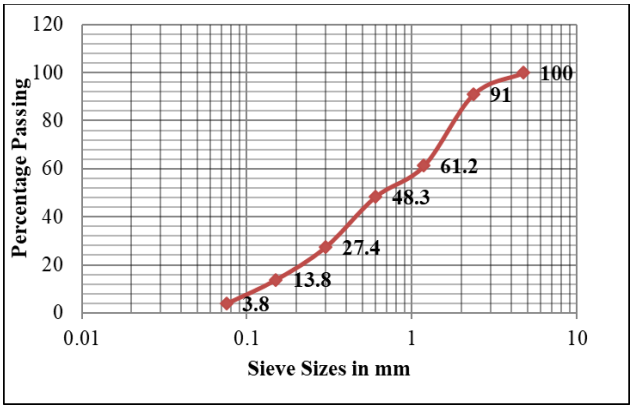
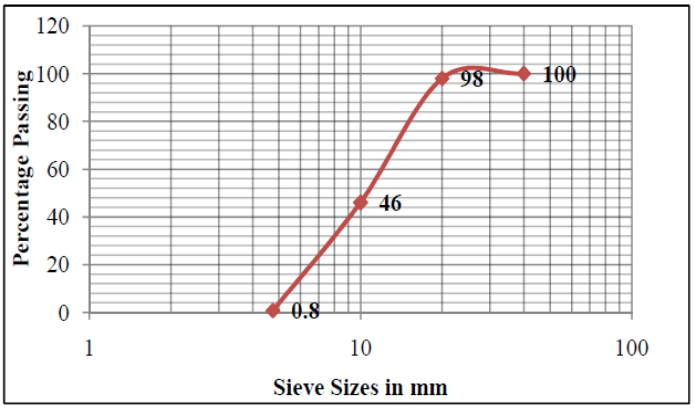


Figure 4: Sieve analysis of M Sand



5.3. Impact Test Results

Figure 10 shows the variation in impact test results of the concrete specimen prepared by replacing natural sand by M-sand at the percentages of 0%, 45%, 50%, 55% and 100%. Results shows that energy consumption in 55% and 100% replacement

shows increase in energy consumption that is about 47.99% and 40% for initial crack, 48.38% and 38.7% for ultimate failure when compare to 0% replacement. It can be concluded that 100% and 55% substitution of river sand by M sand has a threshold values for an optimal performance in energy adsorption and crack resistance.

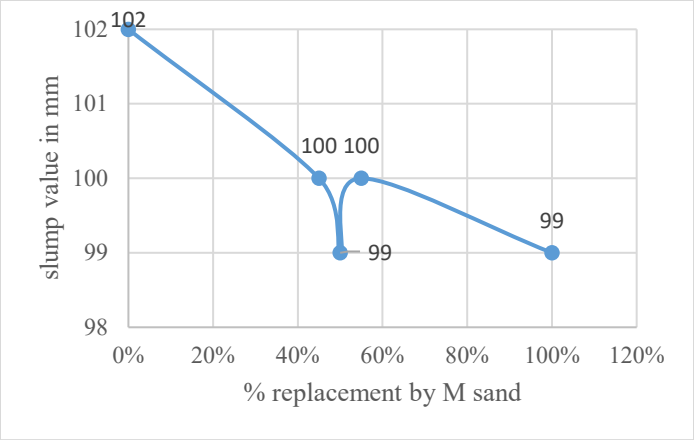


Figure 6: Slump test graph



Figure 6: Flexural strength test



Figure 7: Impact test on concrete

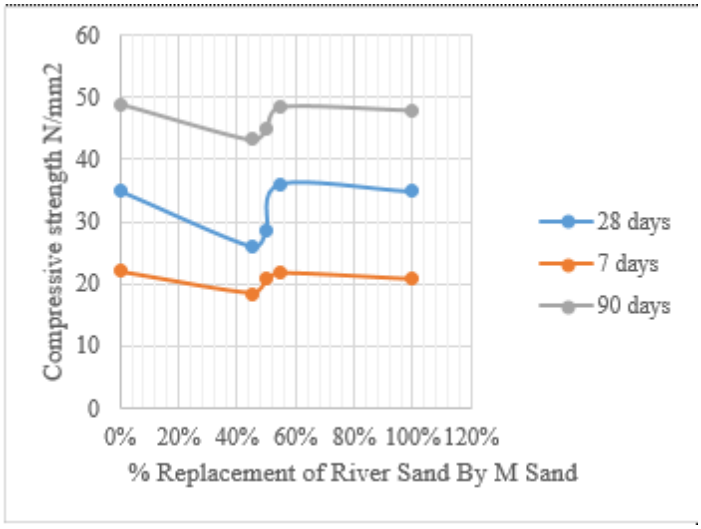
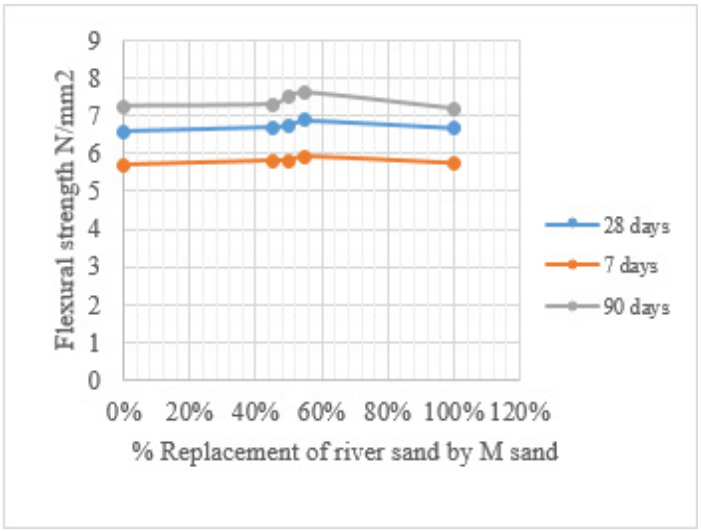


Figure 8: Compressive strength test graph



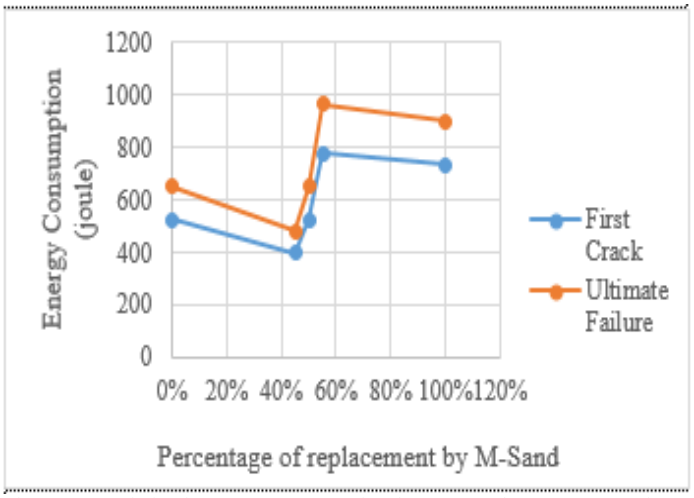


Figure 10: Variation in Energy Consumption for M20 grade Concrete Specimen

Conclusion

M-sand and River sand fineness value are 3 and 3.45 respectively, both falls under zone-2, it shows the M-sand

properties is similar to river sand, M-sand is slightly coarser as compared to river sand.

By 100% and 55% incorporation of M sand by river sand concrete can achieve higher flexural strength and compressive strength of concrete at 7 days, 28 days and 90 days.

Concrete can achieve optimal performance in energy adsorption and crack resistance by 100% and 55% substitution river sand by M-sand.

An incorporation of 100% and 55% of M sand can be advised to use as fine aggregate to enhance the strength of Concrete.

Acknowledgement

The authors would like to acknowledge financial support from iRMC Bold2025, Universiti Tenaga Nasional, Malaysia.Grant Code: RJO 1043 6494.

References

[1] Altamashuddunkhan, N., & Siti Alliyah, M, "The Effect of Replacement of Natural Sand by Manufactured Sand on the Properties of the Concrete. 2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS)", 22–23, 2018. DOI: 10.1109/ICETAS.2018.8629206

[2] A. Zimar, G. Samarawickrama, W. Karunarathna, and S. Jayakody, "Effect of Manufactured Sand as a Replacement for Fine Aggregates in Concrete," 8th Int. Conf. Struct. Constr. Manag., no. October 2017.

[3] P. A. Jadhav and D. K. Kulkarni, "Effect of replacement of natural sand by manufactured sand on the properties of cement mortar," vol. 3, no. 3, pp. 621–628, 2013.

[4] G. Nithyambigai and F. Ash, "Partial Replacement of Manufactured Sand and Fly Ash in Concrete," vol. 5, no. 6, 2015.

[5] K. U. Mani, N. Sathya, and R. Sakthivel, "International Journal of Modern Trends in Engineering and Research Effect of replacement of River sand by M-sand in high strength concrete," 2014.

[6] R. N. U. Shanmugapriya T., "Optimization of Partial Replacement of M-Sand By Natural Sand in High Performance Concrete With Silica Fume," Int. J. Eng. Sci. Emerg. Technol., vol. 2, no. July, pp. 73–80, 2012.

[7] N. R. Monika and S. Sarankokila, "Partial Replacement of M-Sand on Concrete by Self Curing and Standard Curing Method," pp. 12162–12169, 2015. DOI:10.15680/IJIRSET.2015.0412087

[8] T.Subramani & K.S.Ramesh, " Experimental Study on Partial Replacement of Cement With Fly Ash And Complete Replacement of Sand With M sand," Int. J. Appl. or Innov. Eng. Manag., vol. 4, no. 5, pp. 313–322, 2015.

[9] A. Jayaraman, "Experimental Study on Partial Replacement of Natural Sand with M- Sand and Cement with Lime Stone Powder," vol. 6, no. 2, pp. 948–954, 2014.

[10] T. F. Kala, P. Asha, and S. Elavenil, "Properties of Concrete using Manufactured Sand as Fine Aggregate," Int. J. ChemTech Res., vol. 11, no. 03, pp. 94–100, 2018. DOI: http://dx.doi.org/10.20902/IJCTR.2018.110337

[11] M. Meghashree, "Comparison of Physical Properties between Natural Sand and Manufactured Sand," vol. 3, no. 07, pp. 92–96, 2016.

[12] S. Suresh and J. Revathi, "Effect of M-Sand on Setting Time of High Performance Concrete," Asian J. Res. Soc. Sci. Humanit., vol. 6, no. 10, p. 1648, 2016. DOI:10.5958/2249-7315.2016.01118.7

[13] P. Jaishankar and K. Saravana Raja Mohan, "Experimental investigation on nano particles in high performance concrete," Int. J. ChemTech Res., vol. 8, no. 4, pp. 1666–1670, 2015.

[14] Y. Boopathi and J.Doraikkannan, "Study on M-Sand as a Partial Replacement of Fine Aggregate In Concrete," Int. J. Adv. Res. Trends Eng. Technol. vol. 3, no. 2, pp. 746–749, 2016.

[15] B. K. Meisuh, C. K. Kankam, and T. K. Buabin, "Effect of quarry rock dust on the flexural strength of concrete," Case Stud. Constr. Mater., vol. 8, no. July 2017, pp. 16–22, 2018. DOI: https://doi.org/10.1016/j.cscm.2017.12.002

[16] J. R. Sagura R, "Experimental Study on Mechanical Properties of M-Sand Concrete by Different Curing Methods.," IOSR J. Mech. Civ. Eng., vol. e-ISSN: 22, pp. 19–25.

[17] M. Idrees, " Utilization of Waste Quarry Dust And Marble Powder in Concrete," no. July, 2019.

[18] B. Bahoria, D. K. Parbat, P. B. Nagarnaik, and U. P. Waghe, "Effect of replacement of natural sand by quarry dust and waste plastic on compressive & split tensile strength of M20 concrete," Procedia Eng., no. March, p. 9, 2013.

[19] P. Nanthagopalan and M. Santhanam, "Fresh and hardened properties of self-compacting concrete produced with manufactured sand," Cem. Concr. Compos., vol. 33, no. 3, pp. 353–358, 2011.
DOI: https://doi.org/10.1016/j.cemconcomp.2010.11.005

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 306-312 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Comparative Study of Safety Leading and Lagging Indicators Measuring Project Safety Performance

Sevar Dilkhaz Salahaddin Neamat*

Department of Mechanical Engineering, University of Zakho, 42002, Zakho, Kurdistan Region - Iraq

ARTICLE INFO

Article history:
Received: 30 August, 2019
Accepted: 18 November, 2019
Online: 12 December, 2019

Keywords:
Construction safety injuries
Leading indicators
Safety measures
Lagging indicators

ABSTRACT

The safety management system is recognized by safety leading and lagging indicators, and their correlation with injury rates. The background on this specific subject is vague in definition, labelling, and indicators measurement. The comparativeness between leading and lagging indicators have been introduced in the constructing safety performance projects evaluation. Safety performance leading indicators are the metrics of the safety method in constructing work. While the lagging indicators associated with safety results. Many suggestions have been considered during the usage and selection of the effective leading indicators. Also, the research results outlined that the leading indicators can be used to discriminate the variances in the safety performance of projects. In this specified research, leading and lagging indicators have been reviewed and investigated. Eighteen papers have been investigated from the period of 2010 to 2019 in order to recognize the common leading and lagging indicators. In addition, this review will recognize the gaps in leading and lagging research in order to concentrate on extra studies in that field. Four papers mentioned the correlation between lagging and leading indicators. A scoping review is focusing on the points and the significant ideas associated with the research area in accordance with the history time.

1. Introduction

In construction projects, fatality and disable injuries are larger than industry average by three times (Center for Construction Research and Training 2008). According to [1]–[3] the proactive safety management system have an effective and positive influence on the performance. In the past research many effective and proposed cause have been investigated. An examples of designing for safety management [3]–[5] and schedule-based safety management are mentioned [6], [7]. Nevertheless, during the construction work the safety proactive strategies should occur in order to catch the world-class performance. For example, pre-task preparation, stop work authority, and hazard acknowledgement programs used to control and recognize possible hazards which previously resulted in injuries [1], [2], [8]–[10].

To recognize the main reasons around site results of incidents and injuries, the leading indicators are identified. In addition, they have unrestrained hazards and administrative policies together with practices, agendas, and applies that monitor, control, and remove these hazards. The leading indicators are obtainable in daily work of construction organization and have a better role in measuring the safety performance in the workplace.

In this study, a wide range of various leading indicators been investigated, including Alcohol / Drug Testing, Attitudes and Safety Climate, Fall Protection… until, Training and Job Safety Talks Worker.

The introduction of the comparison between using leading indicators of the safety performance correlatively with lagging indicators been proposed. Within the constructing projects the term of leading indicators used to assess the safety process, meanwhile lagging indicators describe the safety outcomes. This study investigates the most common list of leading and lagging factors associated with the selected articles. An example of lagging indicators, Behavior first aid and Reported Incidents. The research contains detailed overview of safety leading and lagging indicators and suggests a dissimilarity between leading indicators and lagging indicators of safety expectation. Researchers can use the results to explain and mark resource expenditures by means of universal scientific indication. Some other studies such as [10]–[14] used different technologies and others used web technology tools as in [15]–[20] to perform investigations.

* Sevar Dilkhaz Salahaddin Neamat, Email: sevar.dilkhaz@uoz.edu.krd

These positive approaches are seldom officially dignified or checked to recruit positive replies when results are not seen. Investigation was wanted to improved apprehend the strategies that might help as forecasters of safety performance and how they can monitor, and control safety risk. This study aims in determining what relationships are found when contractor administrative data is used to measure safety performance.

18 recent articles had been reviewed and analyzed to make a confident conclusion about the leading and lagging indicators to distinguish the positive and negative impacts between them and how they correlate to each other's. Only four paper discuss the interrelation between leading and lagging indicators. Ten leading indicators and two lagging indicators had been resulted after the analyzing of the 19 indicator to be further investigated in the future works.

2. Methodology of the Literature Review

In the beginning, it is a compulsion to cultivate a list of factors in order to choose which administrative data must be selected. To develop a widespread list of leading with lagging safety indicators that have been used in construction projects, a scoping review was accomplished. Additionally, to add a wide-ranging list of indicators, this study sums up the relationship between leading and lagging indicators in previous researches. In order to choose the suitable indicators for the review study, the contents of the earlier studies have been depended. Appropriate less1ons were chosen using the electronic database such as PubMed, Scopus and Web of Science. The information coverage is from both medical and engineering research fields, very crucial in safety construction subjects.

Fifty papers were chosen by means of the search plan. Afterward the duplicate articles were omitted, 48 researches were reread for enclosure depended on the following criteria: 1) Construction fields association; 2) safety factors; 3) recognizing leading and lagging indicators; and 4) academic journals studies.

After that, 38 of the 48 articles remained with 10 exclusions due to duplications and not related to the desirable subject. Later on, 20 articles underwent further examination. Only articles that assessed the safety indicators at a construction project class were included. A total of 2 papers were unenclosed. Finally, the articles that have a direct relation to construction and safety indicators were involved. The remaining articles that have been depended in this study were 18 articles. We prepared some excel sheets to track the researchers, publication years, research types, and the goals together with the indicators that have been used with leading and lagging group. The steps for searching process are in shown figure 1.

Largely, the papers embrace case studies, pilot studies, and literature reviews. Additionally, only a few researchers participated in the research. Table 1 summarizes the type and goal of the reviewed researches.

As it can clearly be seen from the Table 1, the research goals of the final eighteen articles are wide-ranging. Leading indicators used in constructing work should firstly been developed. Secondly, tested in a construction project setting [1]. Numerous of these papers were focused on the first step – development of indicators. This includes defining [21], [22], developing [23]–[25], or measuring the indicators [26], [29].

The followed articles focused on progressing the indicators such as [21]–[27]. Only a few articles concentrated on the second step which is the testing on construction projects through completion of validation testing of the leading on the lagging indicators [1], [28].

The articles in this review paper were depended on the classified collection of safety indicators. The study categorized the indicators to either leading or lagging indicators. So, fifteen leading and four lagging indicators were comprised. By taking in consideration that not all authors used the same vocabularies for the similar indicators were grouped together for the purpose of this

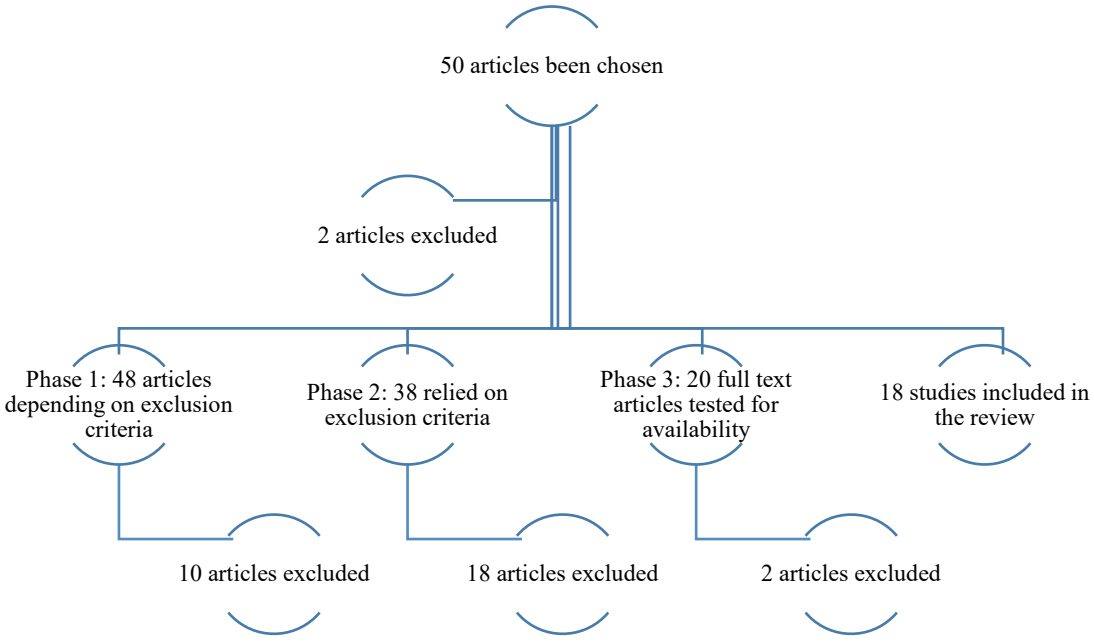


Figure. 1. The searching steps.

Table 1: Summery of Research Types and its Goals

Research	Year	Research Type and Goal
[26][27]	2010	Case Study: to measure the impact of a 5-S assessment tool and safety leading indicators on project safety
[23][24]	2011	Case Study: a once-a-month weighted safety key score and safety climate review were used to assess the safety performance on a construction project
[27][28]	2012	Case Study: analysis of use of a 5-S assessment tool on one construction project to observe how safety leading indicators were reported over the course of the project
[29][30]	2012	Academic framework for using safety performance pointers in safety–critical societies had been presented. It combines 3 sorts of safety performance indicators – outcome, monitor and drive indicators
[30][31]	2013	Review Article: Discusses the difference between passive and active leading indicators within construction safety performance
[31][32]	2013	Case Study: Analysis of leading and lagging indicators of one project to estimate leading indicators below real project circumstances
[32][33]	2014	Review Article: The paper emerged several articles, models utilized in designing of meters, object of study, objectives of interpretation, terms of implementation and revision, actors impacted by the product, and expressions to certify the scientific application and working feasibility
[33][34]	2015	Pilot Study: This research developed a leading indicator-based incentive program for construction projects
[24][25]	2015	Theoretical Article: This article works to create a conceptual framework for developing safety leading indicators for use within the construction industry
[22][23]	2016	Review Article: Discusses the literature on the leading indicator, safety climate, including on a project level.
[28][29]	2017	Case Study: Analysis of the temporal relationships of leading indicators within one five-year construction project.
[25][26]	2017	Focus Groups: This research works to develop indicators that influence site safety that can be measured using safety climate tools
[34][35]	2018	The paper displays a ML method in improving leading indicators that categorize places in accordance to their safety risk in construction projects.
[35][36]	2018	It would be necessary to improve compound key connecting the main features of safety containing the cultural and climatic issues to add a further illustrative picture of platforms' safety performance
[36][37]	2018	Quasi-experimental longitudinal investigation design will be used within 2 Ontario critical clinics. The first stage of the investigation will be based on evaluating the existing OHSMSs consuming the leading indicators. 2nd will pilot test and assess the tailored intervention
[37][38]	2019	In a cross-sectional analysis of CSAP (Contractor safety assessment program) database, several organizational leading safety indicators have been identified, that are associated with safety performance in the construction industry
[38][39]	2019	This research symbolizes safety prequalification reviews presently in usage in the construction industry to recognize methods that contain leading indicators of employee safety performance
[39][40]	2019	This study contains a comprehensive review of safety leading pointer research, suggests a dissimilarity between leading indicators and other approaches of safety prediction, and describes a clear method for distinctive between active and passive indicators.

scoping review. The furthermost communal pointers that were classified as leading indicators were, attitudes and safety climate, site inspections/audits, training and safety talks, and worker safety behavior.

The most popular factors among lagging groups were first aid injuries and lost time injuries. More information can be noticed in Table 2.

Only the indicators that been used in two or more than two studies have been selected in the results. Some indicators were only used by [26], [27]. Though, these writers did not express their indicators or categorize what was measured, beyond a wide classification, such as fall protection. Consequently, some usable indicators by these researches were not further explained. Including, fall protection, ladders and stairs, Personal Protection Equipment (PPE), and railings and covers.

2.1. Results of Comprehensive Literature Review

Ten leading indicators and four lagging indicators were identified over a scoping review. Also, they have been used in researches that concentrated on contractor project data. The lagging and leading indicators will be more clarified below.

2.1.1. Summary of Leading Indicators

In this section, the leading indicators are widely discussed and proposed than lagging indicators. These studies appeared to be concentrated more on the attitude and behavior of the leading indicators such as safety climate, in comparison to avoidance indicators such as site inspections. The familiar leading indicators are proposed below:

Attitudes and Safety Climate: Numerous articles largely premeditated the attitude of employees or managers in the place of work. One specific measure is safety climate. Safety climate is largely well-defined as the employee insights of safety in the job site, specifically as it related to manager opinions [22], [40].

Housekeeping: Housekeeping is considered a physical hazard indicator in the workplace [25]. Housekeeping is a comprehensive phrase restricting the cleanliness of a site. It comprises features such as waste removal, and material and equipment storage (Construction Projects, 1991).

Near Miss: Is deliberated a close call incident. It may cause injuries or totally damaged, but no damage was caused [1] as the navigation of how to manage and control the hazards and avoid it is occurrence in the workplace. Through using PPE, change in the task sequence and controlling of guards [49].

Safety Corrections: In the [26], [27] Safety Corrections are measured as safety correction frequencies prepared by workers on an ad hoc basis. Conversely, [23], measured non-compliance measurements during the site inspections.

Site Inspections: Site inspections, and their similar coordinate, site examining, are tests that evaluate the hazards number on the workplace. In [1] used a scoring system where opinions were subtracted depended on the number of safety violations that happened on the site. Site checks are checklists prepared by some familiar individuals such as a supervisor or safety representative [28].

Subcontractor Safety: Subcontractor safety can be noticed at various levels. [24] Focused on the subcontractor managing of safety as an indicator in the advance safety condition. Also, in project construction until now they did not offer a measurement methodology. In [33] the calculated subcontractor safety performance system depended on the subjective calculation which investigates safe and unsafe work comments to add a single score. [30] Added examples on how subcontractor safety system can be worked and restrained subcontractor safety score not from job work actions Prior to subcontractor award relied on subcontractor safety history, safety policy and their site-specific safety program and pre-task safety plan [24].

Table 2: Leading and Lagging Indicators

Research	*Year*	*Leading Indicators*															*Lagging Indicators*			
		Alcohol / Drug Testing	*Attitudes and Safety Climate*	*Fall Protection*	*Housekeeping*	*Ladders and Stairs*	*Near Miss*	*Pre-Task Safety Plans*	*Personal Protection Equipment*	*Railings and Covers*	*Safety Corrections*	*Safety Positive Reinforcement*	*Site Inspections / Audits*	*Subcontractor Safety*	*Training / Job Safety Talks*	*Worker Safety Behavior*	*First Aid Injuries*	*Lost Time Injuries*	*Members of the Public Injured*	*Reported Incidents*
[26][27]	*2010*			✓	✓	✓			✓	✓	✓									
[23][24]	*2011*		✓								✓		✓		✓	✓	✓	✓	✓	
[41][42]	*2012*			✓	✓	✓			✓	✓	✓									
[29][30]	*2012*		✓				✓	✓			✓	✓	✓	✓			✓	✓		
[30][31]	*2013*		✓				✓					✓	✓	✓	✓					
[1]	*2013*						✓	✓					✓			✓	✓	✓		✓
[32][33]	*2014*	✓	✓	✓	✓				✓		✓								✓	
[24][25]	*2015*		✓		✓			✓						✓	✓	✓	✓	✓		✓
[33][34]	*2015*												✓	✓	✓					
[22][23]	*2016*		✓													✓				
[28][29]	*2017*	✓						✓			✓		✓		✓			✓		
[25][26]	*2017*		✓												✓	✓				
[34][35]	*2018*		✓				✓						✓						✓	✓
[36][37]	*2018*															✓	✓	✓	✓	✓
[35][36]	*2018*		✓																	
[37][38]	*2019*	✓		✓	✓	✓	✓		✓			✓					✓	✓		
[38][39]	*2019*						✓					✓	✓	✓	✓	✓				

[39][40]	*2019*		✓					✓					✓		✓	✓				✓

Training/Job Safety Talks: [24] point out safety discussions, safety instruction, and safety training for supervisors to edit safety conditions considered as safety practices. Training has often measured a side of safety climate [22], [25]. Toolbox talks, or job safety talks, used to change the conversation between the management and workers on a steady base and can be seen as a type of safety communication [28].

Worker Safety Behavior: Worker safety behavior was dignified in lots of studies. It includes measures such as worker participation in safety [22], employee safety inspiration, and worker safety capability [24]. [1] depended on using checklists instead of safety professional evaluation to evaluate workers' safety behavior.

As shown by the leading indicators explained above, the leading indicators focused on hazard identification or correction such as site inspections or training, as well as attitudes and behaviors such as safety climate. All indicators were related to behaviors or attitudes that promote accident prevention rather than negative safety outcome.

2.1.2. Summary of Lagging Indicators

Lagging indicators, have been researched within a longer time than leading, are typically easily accessible and well documented. Yet, lagging indicators were less commonly used in the studies of this specified review. Lagging indicators are generally accessible as they are reporting measures that are required by various health and safety authorities. It used to track company performance and to monitor injury trends (OHSA, 1990). It makes the research on lagging indicators convenient. Yet, despite the convenience, there seems to be a shift towards leading indicators in the current research. This shift leaves common lagging indicators such as first aid injuries and lost time injuries, less used in research.

There are many reasons as to why lagging indicators are less commonly researched. One is the focus on injury prevention in research and in occupational health and safety practice. Researchers and practitioners have chosen to focus on indicators that can lead to injury prevention, such as site inspections, instead of indicators that document safety outcomes, such as injuries. The thought is that leading indicators allow for research to be more proactive, rather than retroactive with lagging indicators

First Aid Injuries: First aid injuries include any injury that can be treated with minimal first aid and require no further medical treatment [1]. First aid injuries are recorded on site as part of the reporting requirements in the OHSA (1990). This reporting requirement is similar across many jurisdictions. This measure was included by [23] as well.

Lost Time Injuries: Lost Time Injuries or Total Recordable Injury Rate is the ratio used by the United States Occupational Safety and Health Administration (OSHA) including any injuries that require medical treatment beyond first aid [1].

A subset of the Total Recordable Injury Rate is the Lost Time Injuries, where workers miss work leading to loss of earnings (Ontario Health and Safety Act, 1990). The documentation of lagging indicators is related to safety authority documentation requirements. The type of indicators used, whether total recordable injury rate or lost time injuries, is often based on the safety authority in the research area. Lost time injury is the more commonly used metric, as it is used by the Workplace Safety and Insurance Board (WSIB) under the reporting requirements of the OHSA (Ontario Health and Safety Act, 1990).

As it can be seen through the factors recognized in this study, the lagging indicators were correlated to undesirable safety results and precisely varying levels of injuries. Lagging indicators are willingly obtainable depended on government recording requirements, but are not widely used in the investigation in favors of leading indicators.

2.1.3. Interaction of Leading and Lagging Indicators

The scoping review concentrated on how leading and lagging indicators are correlated to one another during construction projects. To the best of our knowledge, only four studies were found investigating such studies. Therefore, this interaction became the main part of the project discussion. Unfortunately, although these four studies added some research into leading and lagging indicators in construction projects, the existence of limited articles being obtainable, presented the necessity for more research in this field.

Research [28] had been accomplished the analyses of leading indicators temporally. The researchers are not insured if the leading indicators are really leading, and how time-consuming until the implementation of leading indicators resulted in the improvements of the lagging indicators. They planned 11 leading indicators counting safety talks, hazards testified and safety inspections. These academics used frequency data gathered by sub-contractors and workers on a huge, five-year construction project in Australia.

This incidence data experienced numerous stages of pre-processing such as modification for man hours previous to being involved in results. This research found that the 'leading' feature of leading indicators is really energetic in nature. Leading indicators did not develop the lagging indicator, precisely the full recordable injury rate is estimated in the foreseeable technique.

This led the researchers to consider that the safety leading indicators is complicated. Additionally, they suggest that there is a two-directional correlation between leading and lagging meters. Similar, this paper was not supporting the considerable temporal feature of leading indicators. They resulted through the searching that leading indicators both led and lagged. For instance, toolbox talks caused to minimize in injuries for the first 4 months then damages led toolbox talks for the next 2 months. Herein matter, toolbox talks and injuries considered as a leading and lagging indicator, and were consequently bidirectional. This demonstrates that theory that forms the basis leading and lagging indicators is very simple in relation to the nature of construction complexity. These writers recommend concentrating fewer on leading and lagging categories, and somewhat tag indicators as positive and negative. Whereas the research almost provided the research on leading and lagging indicators, they are unsuccessful in verifying in what way the importance of their research effects management

and construction companies. In spite of that, they pay attention to the significance of this paper on other research. Such that, if the use of leading and lagging expressions possibly will outcome with struggles in the research setting, would it not also cause complications in management interference? This specific question and many still in vague.

The authors of [49] evaluated the active and passive leading indicators relationship in the safety construction projects by using meta-analysis technique. Depending on [42] and [43] , the analyzing been prepared. The meta-analysis process consisted of four stages, literature reviewing, coding studies, normalizing effect sizes, computing the combined effect size. The researchers used the followed search engines to collect the papers ASCE, Web of Science, Scopus, Engineering Village, and Google Scholar. Moreover, many different keywords been used to search such as, safety management system, safety program, construction safety practices, safety performance, safety strategies, safety leading indicators, and proactive indicators. 114 papers had been selected to take apart in the study. Some conditions must be present in the studies that participate in the meta-analysis procedure, if they: 1) examined the association between either active or passive safety leading indicators and accidents or injury; 2) re-counted the effect size (e.g., correlation values) or sufficient info to compute the effect size; 3) tested data from the construction industry; and 4) were peer studied.

As a result, the analyzing software found that the effect sizes of correlation of leading and lagging are extensively different, nine construction safety leading indicators participated in this analysis. The effect sizes of the relationships among safety inspection and observation and injury (r = 0.51; 95%CI = 0.30–0.67) and between pre-task safety meeting and injury were very large (r = 0.45; 95%CI = 0.32–0.57). For the nine passive leading indicators, eight were significant ($p < 0.05$), the relationship between injury rate and safety record (r =0.56; 95%CI = 0.20–0.79) and safety resources (r =0.48; 95% CI = 0.28 to 0.65) had large effect sizes. Staffing for safety (r = 0.44; 95%CI =0.12–0.68), owner involvement (r =0.45; 95%CI =0.16–0.67), training and orientation (r = 0.42; 95%CI = 0.10–0.66), personal protective equipment (r = 0.40; 95%CI =0.17–0.58), and incentives programs (r = 0.30; 95%CI =0.15–0.43) were all moderate. Finally, the effect size of safety inspections and observation was low (r = 0.27; 95%CI = 0.12–0.41), and that of pre-task meetings was not significant (p = 0.103). Finally, the results ensured that the pre-task safety meetings considered an important forecaster for future performance measurement. If dignified and preserved as an active leading indicator. However, considering pre-task safety meetings as a passive pointer is not prognostic.

In [34], the research demonstrates a method in order to improve the leading indicators that categorize the location according to construction work safety risk. The key forms of statistics involved is the safety examination accounts, accident cases and project-related data. The data achieved from a big contractor company in Singapore and collected from 2010 till 2016. From 33 independent variables, 13 input variables had been chosen by a connection of Boruta feature collection technique and decision tree. Six of 13 input variables are project-related (type of project, project possession, contract amount, percent finished, level of postponement and plan manpower) and 7 of them are matters in the contractor's safety inspection checklists (crane/lifting operations, scaffold, mechanical-elevated working platform, falling hazards/openings, environmental management, good practices and weighted safety inspection score). The author used 5 common machine learning algorithms to run models for the forecast of accident existence and harshness. Throughout predicting the validity, random forest (RF) added the greatest forecast performance with an accurateness of 0.78 and has attained a considerable power of arrangement with Weighted-Kappa Statistics of 0.70. Associating with parallel studies, this outcome is talented. The forecast (i.e. the output variable) added by the RF model can be used as a safety leading pointer of the risk level of a place. It is suggested that the prognostic RF model be organized in construction administrations, particularly large public and private creators, contractors and industry relations, to offer a monthly prediction of project safety performance so that preventive examinations and involvements can be fulfilled in a further directed manner.

In [1] the author gathered both leading and lagging indicators for a single constructing project in order to reveal any relation between the indicators of leading and lagging. The three leading indicators recycled were pre-task plan review, worker safe behavior remark mark, and site safety audit score. After that, they have been investigated to show a correlation between the four lagging indicators: first aid or emergency treatment, near-miss incidents, OSHA recordable incidents, and all project incidents. This info was gathered for 37 weeks by safety professionals. The researchers found the associations between pre-task plan review and total incidents (r =-0.507), pre-task plan review and first aid (emergency treatment) (r =-0.573), worker safe behavior observations and whole incidents (r =-0.588), worker safe behavior observations and first aid (r =-0.635). Although the relationships observed were talented, the statistical significance for the correlation coefficients were not added. Additionally, the research based on optimistic results to clarify that out of 12 only 4 only correlations sustained the research hypothesis. The writer added little dialogue of the undesirable results or afford results that did not support their 19 hypotheses. Meanwhile, this research paper started the conversation on leading and lagging indicators on construction projects, an additional vigorous investigation should be done.

The four studied that discussed above added the information in leading and lagging during a construction job. The pilot/case study type of study delivered the vision into the actual relations of these indicators, but superior examples want to be recycled in extra research. Moreover, the results showed that the leading and lagging are not easily correlated and are complex in construction work. While there is a lack of info on the correlation between leading and lagging, some decisions must be done depending on the articles examined. Lagging indicators based on injuries and incidents. All lagging pointers had negative results, corresponding the description of lagging indicators. Further hand, leading indicators can be classified into categories: attitudes and avoidance actions. Instances of leading indicators that evaluate outlooks contain the safety climate and manager attitude towards safety. Illustrations of leading indicators concentrated on the protective measures such as, employee safe behaviors, site safety inspections and pre-task safety plans. Leading indicators depended on the popular attitudes, this result is because the avoidance leading

indicators measurement not always negatively related to lagging indicators such as pre-task plan review and site inspections, shown by researches [1] and [23].

3. Conclusion

In conclusion, the sample size of 18 recent articles have been reviewed and analyzed to make a confident conclusion about the leading and lagging indicators in construction. The research is made in order to distinguish the positive and negative impacts between leading and lagging indicators and how they correlate to each other's. Only four papers discuss the interrelation between leading and lagging indicators. Ten leading indicators and two lagging indicators have been resulted after analyzing 19 indicators to be further investigated in the future works, which focused on contractor project data. This review paper will be very benefit to be a start point in the deeply future research.

References

[1] J. Wanberg, C. Harper, M. R. Hallowell, and S. Rajendran, "Relationship between construction safety and quality performance," J. Constr. Eng. Manag., **139(**10), p. 04013003, 2013.

[2] M. R. Hallowell and J. A. Gambatese, "Construction safety risk mitigation," J. Constr. Eng. Manag., **135(**12), 1316–1323, 2009.

[3] S. Neamat, "Factors Affecting Project Performance in Kurdistan Region of Iraq," Int. J. Adv. Eng. Res. Sci., **4(**5).

[4] J. A. Gambatese, J. W. Hinze, and C. T. Haas, "Tool to design for construction worker safety," J. Archit. Eng., **3**(1), 32–41, 1997.

[5] T. M. Toole, "Construction site safety roles," J. Constr. Eng. Manag., **128**(3), 203–210, 2002.

[6] N. A. Kartam, "Integrating safety and health performance into construction CPM," J. Constr. Eng. Manag., **123**(2), 121–126, 1997.

[7] J. Hinze, J. Nelson, and R. Evans, "Software integration of safety in construction schedules," presented at the Proceedings of the 4th Triennial International Conference, Rethinking and Revitalizing Construction Safety, Health, Environment and Quality, Port Elizabeth, South, Africa, 2005, 16–17.

[8] P. X. Zou, "Fostering a strong construction safety culture," Leadersh. Manag. Eng., **11**(1), 11–22, 2010.

[9] K. Jacksi, "Database Teaching in Different Universities: A Phenomenographic Research," Int. J. Emerg. Technol. Comput. Appl. Sci., **2**(12), 96–100, May 2015.

[10] S. Neamat and I. Yitmen, "Factors Affecting the Innovation and Competitiveness in Kurdistan Region of Iraq Construction Industry," Int. J. Adv. Eng. Res. Sci. IJAERS, **4**(2), 157–162, 2017.

[11] S. Neamat, "The Development Of Management Control Systems Framework In Public-Private Partnerships," Int. J. Sci. Technol. Res., **8** (10),299–305, 2019.

[12] S. Neamat, "A Developed Framework for Energy Technology Sustainability Assessment," Int. J. Innov. Technol. Explor. Eng. IJITEE, **9** (1), 832–838, 2019.

[13] S. Neamat, "Models Developed for Creep of High Strength Concrete," Infogain Publ, **3** (3), 174–80, 2017.

[14] S. D. Salahaddin, "Factors Affecting the Competitiveness and Innovation in Northern Iraq Construction Industry," 2016.

[15] K. Jacksi and S. M. Abass, "Development History of the World Wide Web."

[16] K. Jacksi, "Design and Implementation of E-Campus Ontology with a Hybrid Software Engineering Methodology."

[17] K. Jacksi, S. R. M. Zeebaree, and N. Dimililer, "LOD Explorer: Presenting the Web of Data," Int. J. Adv. Comput. Sci. Appl. IJACSA, **9** (1), 2018.

[18] M. A. Sadeeq, S. R. Zeebaree, R. Qashi, S. H. Ahmed, and K. Jacksi, "Internet of Things Security: A Survey," presented at the 2018 International Conference on Advanced Science and Engineering (ICOASE), 2018, 162–166.

[19] K. Jacksi, N. Dimililer, and S. R. M. Zeebaree, "A Survey of Exploratory Search Systems Based on LOD Resources," in PROCEEDINGS OF THE 5TH INTERNATIONAL CONFERENCE ON COMPUTING & INFORMATICS, COLL ARTS & SCI, INFOR TECHNOL BLDG, SINTOK, KEDAH 06010, MALAYSIA, 2015, 501–509.

[20] K. Jacksi, N. Dimililer, and S. R. Zeebaree, "State of the Art Exploration Systems for Linked Data: A Review," Int. J. Adv. Comput. Sci. Appl. IJACSA, **7(**11), 155–164, 2016.

[21] J. Hinze, S. Thurman, and A. Wehle, "Leading indicators of construction safety performance," Saf. Sci., **51**(1), 23–28, 2013.

[22] N. V. Schwatka, S. Hecker, and L. M. Goldenhar, "Defining and measuring safety climate: a review of the construction industry literature," Ann. Occup. Hyg., **60** (5), 537–550, 2016.

[23] H. Lingard, R. Wakefield, and P. Cashin, "The development and testing of a hierarchical measure of project OHS performance," Eng. Constr. Archit. Manag., **18** (1), 30–49, 2011.

[24] B. H. Guo and T. W. Yiu, "Developing leading indicators to monitor the safety conditions of construction projects," J. Manag. Eng., **32**(1), p. 04015016, 2015.

[25] M. Niu, R. M. Leicht, and S. Rowlinson, "Developing safety climate indicators in a construction working environment," Pract. Period. Struct. Des. Constr., **22**(4), p. 04017019, 2017.

[26] K. Ng, A. Laurlund, G. Howell, and G. Lancos, "An experiment with leading indicators for safety," presented at the Proceedings Annual International Group for Lean Construction Conference, 2010, 253–262.

[27] K. Ng, A. Laurlund, G. Howell, and G. Lancos, "lean safety: using leading indicators of safety Incidents to improve construction safety," presented at the Proceedings for the 20th Annual Conference of the International Group for Lean Construction: are We Near a Tipping Point, 2012, p. 173.

[28] H. Lingard, M. Hallowell, R. Salas, and P. Pirzadeh, "Leading or lagging? Temporal analysis of safety indicators on a large infrastructure construction project," Saf. Sci., **91**, 206–220, 2017.

[29] T. Reiman and E. Pietikäinen, "Leading indicators of system safety–monitoring and driving the organizational safety potential," Saf. Sci., vol. **50**(10), 1993–2000, 2012.

[30] J. Hinze, S. Thurman, and A. Wehle, "Leading indicators of construction safety performance," Saf. Sci., **51**(1), 23–28, 2013.

[31] S. Rajendran, "Enhancing construction worker safety performance using leading indicators," Pract. Period. Struct. Des. Constr., **18**(1), 45–51, 2012.

[32] G. Delatour, P. Laclémence, D. Calcei, and C. Mazri, "Safety performance indicators: a questioning diversity," presented at the 6. International Conference on Safety & Environment in Process & Power Industry (CISAP-6), 2014, **36**, 55–60.

[33] E. H. Sparer, R. F. Herrick, and J. T. Dennerlein, "Development of a safety communication and recognition program for construction," New Solut. J. Environ. Occup. Health Policy, **25**(1), 42–58, 2015.

[34] C. Q. Poh, C. U. Ubeynarayana, and Y. M. Goh, "Safety leading indicators for construction sites: A machine learning approach," Autom. Constr., **93**, 375–386, 2018.

[35] K. H. D. Tang, S. Z. M. Dawal, and E. U. Olugu, "A review of the offshore oil and gas safety indices," Saf. Sci., **109,** 344–352, 2018.

[36] J. M. Almost et al., "A study of leading indicators for occupational health and safety management systems in healthcare," BMC Health Serv. Res., **18**(1), p. 296, 2018.

[37] J. Manjourides and J. T. Dennerlein, "Testing the associations between leading and lagging indicators in a contractor safety pre-qualification database," Am. J. Ind. Med., **62**(4), 317–324, 2019.

[38] K.-H. Liu, J. Tessler, L. A. Murphy, C.-C. Chang, and J. T. Dennerlein, "The gap between tools and best practice: an analysis of safety prequalification surveys in the construction industry," NEW Solut. J. Environ. Occup. Health Policy, **28**(4), 683–703, 2019.

[39] W. M. Alruqi and M. R. Hallowell, "Critical Success Factors for Construction Safety: Review and Meta-Analysis of Safety Leading Indicators," J. Constr. Eng. Manag., **145**(3), p. 04019005, 2019.

[40] S. Neamat, "Risk Assessment for Uzun Construction and Real Estate Company in TRNC," Sch. J. Econ. Bus. Manag., **5**(3), 332–343, 2018.

[41] K. Ng, A. Laurlund, G. Howell, and G. Lancos, "lean safety: using leading indicators of safety Incidents to improve construction safety," presented at the Proceedings for the 20th Annual Conference of the International Group for Lean Construction: are We Near a Tipping Point, 2012, p. 173.

[42] N. A. Card, "Applied meta-analysis for social science research: Methodology in the social sciences," N. Y. Guilford, 2011.

[43] J. Tenboer et al., "Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein," Science, vol. **346**(6214), 1242–1246, 2014.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 306-316 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Multimode Control and Simulation of 6-DOF Robotic Arm in ROS

Rajesh Kannan Megalingam*, Raviteja Geesala, Ruthvik Rangaiah Chanda, Nigam Katta

Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, 690525, India

ARTICLE INFO

Article history:
Received: 15 January, 2020
Accepted: 02 May, 2020
Online: 29 May, 2020

Keywords:
Inverse Kinematics
Robot Operating System (ROS)
Graphical User Interface
6 DOF (Degree of freedom)
Control Interfaces.

ABSTRACT

The paper proposes the design and simulation of a 6 Degree of Freedom (DOF) robotic arm, tailored for the coconut crop harvesting, assistive robots like wheelchairs and Home robots, Search and rescue robots for disastrous environments, Collaborative robots for research use. A kinematics-based solution has been developed for the robotic arm which makes it easier to operate and use. Keyboard, GUI, Joystick are the three control interfaces used in the paper. The robotic control interfaces proposed in the paper were developed using the Robot Operating System (ROS). The 6- DOF articulated robotic arm was designed and visualized in RVIZ. The kinematics helped for the easy manipulation of the robotic arm with the end effector. The methodologies proposed in the research work are easy to operate and inexpensive. The designed 6 DOF robotic arm, the first three DOFs are for positioning of the robot's arm, while the residual three are used for manipulation of the gripper.

1. Introduction

The robotic arm is a programmable mechanical arm that works similar to a human arm. Robotic arms played a significant role in the process of industrial automation. The human-like dexterity of these robotic arms makes them efficacious in diverse applications in a variety of industries - manufacturing, atomic power plants, space exploration, material handling, painting, drilling, agriculture deployments [1], assistive robotics applications [2] and numerous other applications. The robotic arm typically comprises an end effector that is designed to manipulate and govern with the surroundings. The 6 DOF robot arm is designed to manipulate and govern with the surroundings. The 6 DOF is to pivot in 6 different ways that mimic a human arm. The major issues concerned in an industrial robotic arm are its mechanical structure and the control mechanism. The control mechanisms can be effectuated by 3 options: keyboard, joystick and slider-based control. Design of a lightweight robotic arm which can be compatible with any kind of robotic system. In the research, all the proffered control mechanisms adopted inverse kinematics, which makes it easier to control. The proposed control mechanisms are compatible with any other complex robotic systems of the same degrees of freedom. The dexterity of the robotic manipulator depends on the degree of freedom.

Kinematic analysis [3] is one of the important steps in the design of the robotic arm. The loop equations of complex robotic systems can be deduced from the kinematic equations. Joint angles regulate the motion of the robotic arm. Computer-Aided Design (CAD) software is the platform to create models with the given set of geometrical parameters. The proposed robotic arm was designed in Solid Works CAD software considering all the given sets of geometrical parameters. Robot Operating System (ROS) provides an integrated platform for the control of robotic systems. ROS is a special kind of framework initially developed with the purpose of working on robots in the research domains. In order to understand how the ROS framework works one should be clear about the concept of communication of messages through topic between nodes. Simulation is one of the ways to optimize the design and improve the control of robotic systems. ROS provides a 3-D visualizer (RVIZ) which helps in the visualization of the pose. With a properly-set URDF (Unified Robot Description Format) file, one can visualize the robot model in RVIZ. RVIZ is the simulation software used for the control of the 6-DOF robotic arm.

The design and development of 7-DOF arms [4] and even successful simulations on 11-DOF [5] arm are done by other authors with valid proofs of classic IK algorithms [6] in the discrete-time domain for robots. Taking those inputs into consideration the 6-DOF robotic arm in the paper was designed and tested with various goal positions using different modes of control. Algorithm developed for the keyboard based control [7] of the 6 DOF robotic arm using ROS which makes the control easier. The design and simulation of this robotic arm which is

*Rajesh Kannan Megalingam, India, +91-9496120900, rajeshkannan@ieee.org

visualized in RVIZ. Simulation of a robot is also possible on the Gazebo platform [8] with the help of ROS_control packages and the generation of the robot conFigureuration is in Gazebo, it uses the URDF file and this is extensive use of a simulation platform for a real-life scenario. Simulation of the actuators in the arm can be done using various simulation software's [9] like Matlab and Simulink [10], [11] but choosing RVIZ is preferable as there are many references to proceed. An open-source robot simulator called USARSim can also be used for both research and education of a robot's general architecture [12]. An inverse kinematics algorithm [13] needs to be developed to achieve the goal positions in simulation over these platforms using the traditional method the DH parameter [11], [14] the inverse kinematics solutions [1], [15] for the arm have been derived. Few other authors used ADAMS [16] to simulate and evaluate the arm design but this paper deals with RVIZ simulation with multimode controls. With the help of ROS platform and MoveIt, the control [17] of the simulated arm was done and the three control modes [18] were tested. The sliding mode controller mechanism can be helpful for regulation of robotic arms with unknown behaviors [19]. Laser range method [20] can also be used for positioning and tracking tasks for a 6-DOF robotic arm. This method helps in the determination of the distance from the camera to the target. The workspace [21] of the arm and the path for reaching the goal position varies on the approach and techniques used while modelling the simulation. There are various other control mechanisms like trajectory tracking [22], master slave control [23], vibration control of a robotic arm with input constraints. Testing in real time doesn't show accurate results due to backlash and electric issues. Considering this problem some authors proposed a self-learning algorithm [24] that uses positioning error after each trail. Simulation gives an idea about the performance of the algorithm, but in case if real time implementation performance changes due to several factors. One such thing is the connectivity and the architecture. Few authors explicitly gave intuition into the master-slave conFigureuration based tele existence concept [25]. Considering all this leads, a 6 DOF robotic arm is simulated and tested by using different control mechanisms.

Contributions of the research is as follows:

- Kinematics based solution for the 6 DOF robotic arm is proposed.
- Validation of three control methodologies of the robotic arm.
- Simulation and evaluation of the 6 DOF robotic arm in RVIZ.

2. Architecture

The paper is primarily concerned about the different modes of control. Figure 1 represents the general architectural diagram which consists of control modes, a processing unit, and a 6 DOF robotic arm. ROS is used as the platform to develop software programs.

Control in testing the 6-DOF robotic arm by simulating it with an IK-based solution in RIVZ. The programs and files for the control modes and processing unit entirely return in CPP.

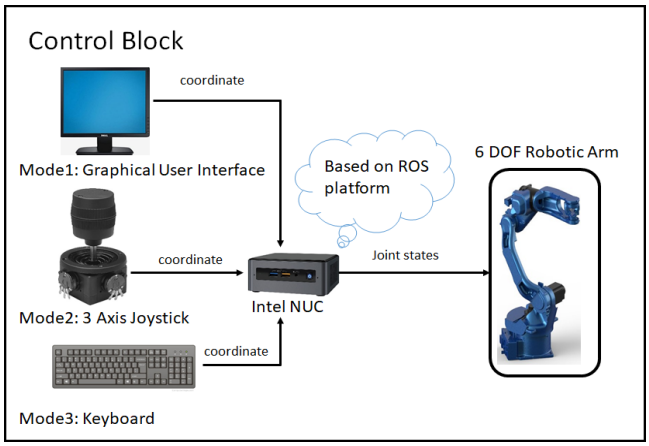


Figure 1: General Architectural diagram.

2.1. Control Block

The control block is the interface between the user and the robotic arm. It consists of three modes of control. The user can control a robotic arm using any of the below modes.

- GUI
- Joystick
- Keyboard

The robotic arms is equipped with a manipulator. A manipulator is a robotic gripper which is used to do dexterity. Figure 2 represents the software architecture of the 6 DOF robotic arm in the ROS platform.

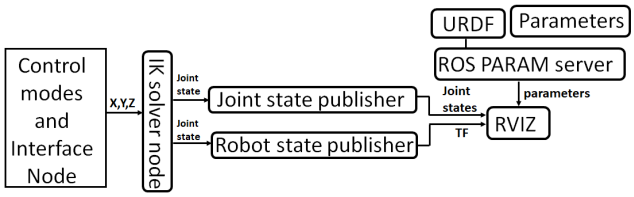


Figure 2: Software Architecture diagram.

The main aim is to control the robotic arm's manipulator position in the RVIZ. Figure. 11, describes an RVIZ axis at the base of the robotic arm. The user can control the manipulator by using any of the control modes. These users input x, y, z coordinates will be published to IK solver node. The IK solver node will compute the joint states which are required for the manipulator to reach to the desired coordinates.

2.1.1 Control modes and Interface nodes

As there are different optional modes to command the robotic arm, interface node acts as an intermediate module that prioritizes the modes. The interface node will act as a bridge between the modes of control and the computing block IK_solver_node. The user can define the order of priority for the modes in the definition file in the interface module. On uploading the entire program, the compiler uses the updated definition file. On simultaneous publishing of data from the three control modes, always the

highest priority mode values will be published to the IK_solver_node from the interface node.

2.1.2. IK Solver node

The Inverse Kinematics equations are generated by using Denavit-Hartenberg. An IK function is written in a computing file where it is passed with x, y, z parameters for computing. And there is a threshold evaluating block of code where the arms reach limit is checked whether it could reach the x, y, z coordinates. If the coordinate is present within the workspace of the arm it generates the individual joint values else, it generates the joint values which head to the maximum point it could reach in the direction heading to the input point. The generated joint values for the desired x, y, z coordinate point will be conveyed to the joint state publisher. Joint state publisher is the one which updates the joint values of the robotic arm simulation.

2.1.3. Graphical User Interface (GUI)

A user-friendly GUI was designed in QT Designer software [26] of version 4.8. Qt Designer is used to create user interface files containing windows and controls. Figure 3 depicts the representation of the graphical user interface which was developed in Qt Designer. The designed window consists of 3 fill in blocks, a SEND COORDINATES toggle button, a HOME toggle button, and a HALT toggle button.

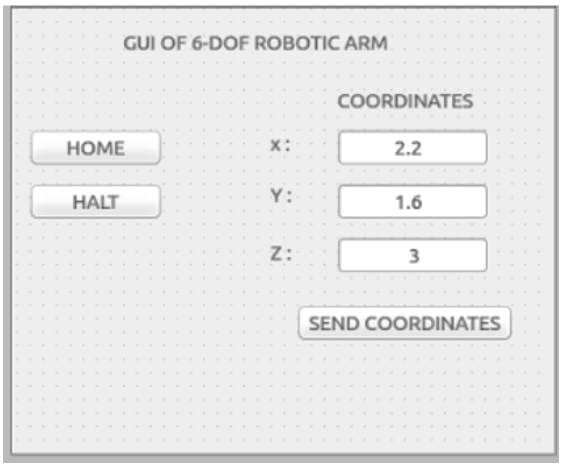


Figure 3: Representation of the graphical user interface.

The user can enter their x y z coordinates to which the user wanted the robotic arm to reach. On clicking the SEND COORDINATE button, the values will get published to the IK solver node. For resetting the robot, a HOME toggle button is designed. Whenever the user presses the toggle button, the robotic arm will reach the predefined pose. A HALT toggle button is also designed which will stop the motion of the robotic arm in case of emergency.

2.1.4. 3- Axis Joystick

The Joystick is another option for controlling the robotic arm. Figure 4 depicts a model of the 3-axes joystick. Arduino is the microcontroller used for taking the joystick values. The joystick will be connected to Arduino analog pins externally; through analog pins the joystick values are fetched.

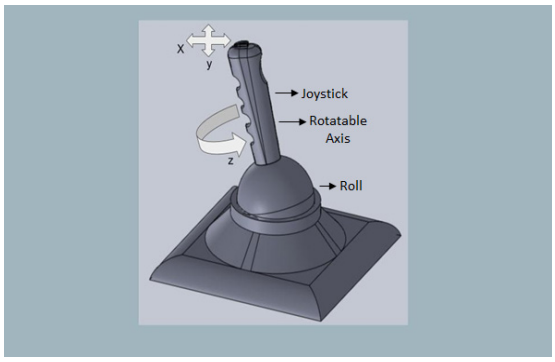


Figure 4: representation of 3 axes joystick

The three axes of the joystick are used to control the x y z coordinates position of the manipulator of 6 DOF robotic arm. If the user moves the joystick's axis in forward direction, x value in the coordinate increases, similarly y and z can also be changed with the other two axes of the joystick. In this way, the user can control the manipulator position.

2.1.5. Keyboard

The keyboard is another mode of control. The keys 'a', 's', 'd', 'z', 'x', 'c', 'h', 'j' are used to control the x, y, z values. When a key from 'a' or 's' or 'd' on the keyboard is pressed, the x, y, z coordinate value increases respectively. When keys from 'z' or 'x' or 'c' are pressed then the coordinate variables x, y, z decreases respectively. 'h' and 'j' are the home pose and halt buttons. These coordinates (x, y, z) are published to the interface node. And from the interface node to the IK solver node for further computation.

3. Design and Implementation

Robotic arms play a vital role in the industrial and social applications, where the design of the arm plays an important role in its function. According to the functionality, there are six main types of industrial robots: Cartesian, SCARA, cylindrical, delta, polar and vertically articulated. The design of the robotic arm that was developed for the research is not compromised in terms of the workspace. Any arm with revolute joints between all its base members falls in the category of the articulated robotic arm. An articulated arm has the ability to rotate all the joints (number of joints is from two to ten). The reason behind the choice for the articulated robotic arm for the research is that it has more work envelope compared to SCARA, Delta, and Cylindrical.

Arm base is called the waist which is vertical to the ground and the upper body of the robot base is connected to the waist through a revolute joint which rotates along the axis of the waist and to which the second DOF (shoulder) is attached as similar to the arm in Figure 5. Third DOF which resembles the elbow of a human arm. The fourth DOF gives yaw movement, fifth gives the pitch and the last DOF gives the roll movement. The 6-DOF articulated robotic arm was designed on Solid works and dimensions are mentioned in the Table 1. The first 3-DOFs give positional (i.e. x, y, and z) coordinates to the end-effector and the last 3-DOFs give the orientation of the end-effector. In the Table 1 base footprint is the base mounting stand for the robotic arm. The arm has 6 links namely base_link, Link1, Link2, Link3, Link4, Link5 and

followed by the gripper. Link1 is the link between joints J1 and J2, Link2 is between joints J2 and J3. Similarly, for Link3, Link4, Link5 are the links with their respective joints.

Assembly file in Solid works is converted into a URDF file using Solid works with a maximum reach length of 1.248 m. Figure 5 and Figure 6 represents the design of the model of a 6-DOF robotic arm in Solid works. URDF is the conventional format for describing robots in ROS. The format is used to simulate a virtual arm. The joint positions, link lengths, and global origin are defined in the Solid works before exporting the assembly file of Solid works into URDF for simulation. The designed arm is simulated and controlled in Rviz using ROS. The evaluation of the designed URDF is done using MoveIt setup assistant which is ROS platform. In this section, the author first loaded the URDF file and created the collision matrix. MoveIt divides the arm joints and manipulators into groups, which helps to have a safer control of the robotic arm by resolving errors due to singularity issues.

Table 1: Joint link lengths.

i	joint	parent	child	Link length "(m)"	Shift in Origin "(m)"
0	Fixed base	Base footprint	base_link	0	0
1	J1	base_link	link1	0.15	0
2	J2	link1	link2	0.14	0
3	J3	link2	link3	0.45	0
4	J4	link3	link4	0.216	0.04
5	J5	link4	link5	0.091	0
6	J6	link5	link6	0.106	0.03
7	gripper	link6	Gripper link	0.095	0
	Total			1.248	

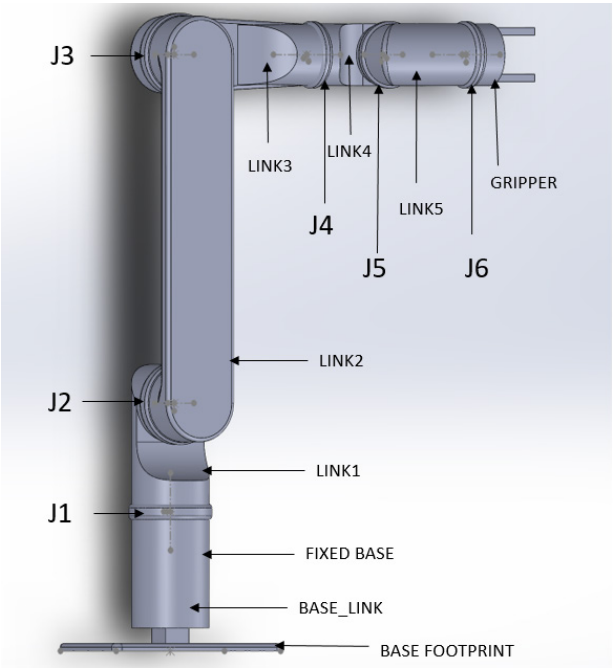


Figure 5. Front view of the Robotic arm

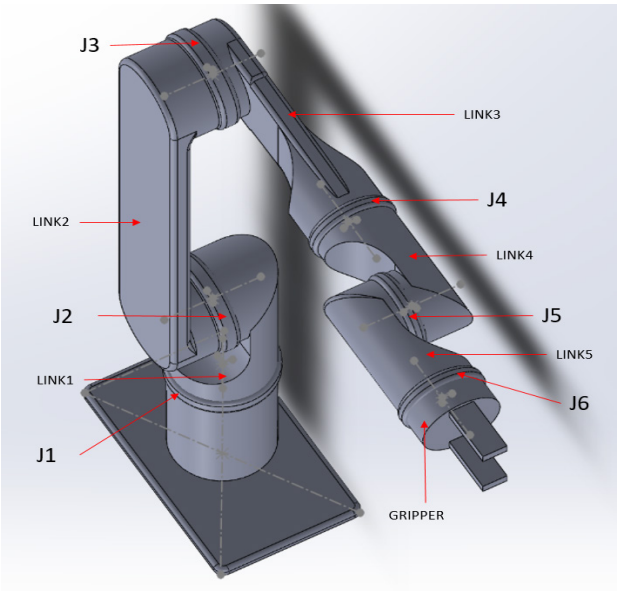


Figure 6. Top view of the Robotic arm

4. Robotic Arm Kinematics

The kinematics of the robotic arm are categorized into forward kinematics and Inverse kinematics. To calculate the kinematics solution the concepts of reference frames, Euler angles and Homogenous transforms are important.

4.1. Reference Frame representation

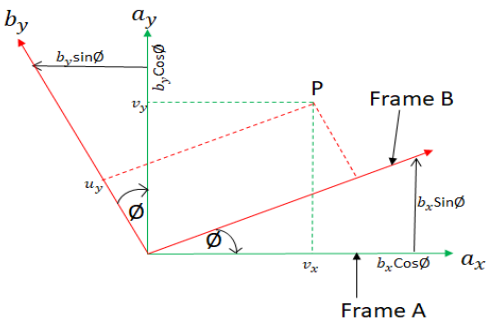
Reference frames play an important role for the kinematics [27] of any arm. The general concept of reference frames is explained with the illustration in Figure 7. Consider point P, in the 2-D Cartesian as shown in Figure 7, is expressed with vector U relative to coordinate frame B. The same point P needs to be represented in frame A with vector V as shown in equations (1), (2) and (3). The matrix representation of the point P in frame A is given in (3). Any point in frame B is multiplied with baR will project it onto frame A.

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \hat{a}_x\hat{b}_x & \hat{a}_x\hat{b}_y \\ \hat{a}_y\hat{b}_y & \hat{a}_y\hat{b}_y \end{bmatrix} \begin{bmatrix} v_x \\ U_y \end{bmatrix} \quad (1)$$

$$\hat{a}_y = \hat{b}_x \sin\phi = \hat{b}_y \cos\phi$$
$$\hat{a}_x = \hat{b}\cos\phi = -\hat{b}_y \sin\phi \quad (2)$$

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \cos\phi & -\sin\phi \\ sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} U_x \\ U_y \end{bmatrix} \quad (3)$$

$${}^{a}_{b}R = \begin{bmatrix} \cos\phi & -\sin\phi \\ sin\phi & \cos\phi \end{bmatrix}$$



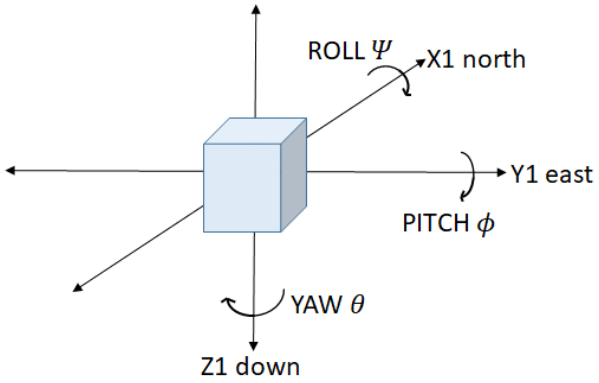
4.2. Euler angles

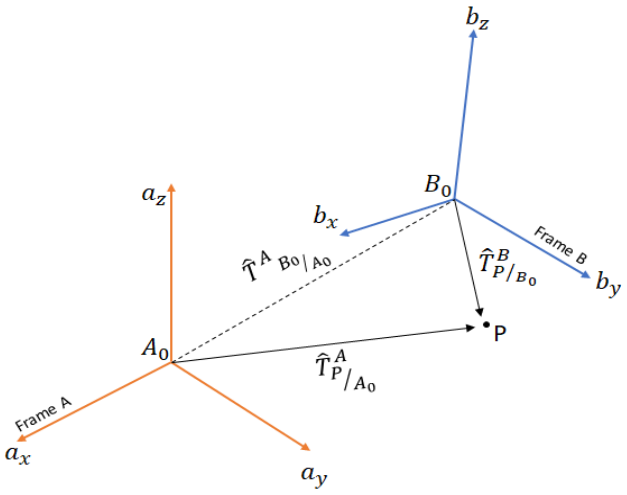
For a rigid body in space as shown in Figure 8 a box is considered as a rigid body, it has 3DOF for rotation those are roll, pitch, and yaw. These are called the Euler angles [28] and the representation of this rotational DOF. Euler angles are a system to describe a sequence or a composition of rotations. Conventionally, the movements about the three axes of rotations and their associated angles are described by the 3D rotation matrices. Equations (4), (5) and (6) show the matrix representation of roll, pitch, and yaw. These are 3D- rotation matrices and anticlockwise in direction.

$$R_x(\Psi) = \begin{bmatrix} cos\,\Psi & -sin\,\Psi & 0 \\ sin\,\Psi & cos\Psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

$$R_y(\phi) = \begin{bmatrix} cos\,\phi & -sin\,\phi & 0 \\ sin\,\phi & cos\,\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$R_z(\theta) = \begin{bmatrix} cos\,\theta & -sin\,\theta & 0 \\ sin\theta & cos\,\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$





4.3. Homogeneous transforms

Homogeneous transform is the representation of the transform matrix of one frame which is both rotated and translated with respect to some other unknown frame. In Figure 9 point in frame, B is transformed onto frame A. And the homogeneous transform is generated using rotational and translational matrices.

Equation (7) the term $\hat{T}^A_{P/A_0}$ is the representation of point P vector with respect to frame A and this is 3*1 matrix. On the other end in the term R^A_B is the rotational matrix of frame B on to frame A and this is a 3*3 matrix. $\hat{T}^A{}_{B_0/A_0}$ Is the translation vector and this is 3*1 matrix. The $\hat{T}^B_{P/B_0}$ is a representation of point P vector with respect to B frame.

$$\begin{bmatrix} \hat{T}^A_{P/A_0} \\ 1 \end{bmatrix} = \begin{bmatrix} R^A_B & \hat{T}^A{}_{B_0/A_0} \\ 0\ \ 0\ \ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{T}^B_{P/B_0} \\ 1 \end{bmatrix} \quad (7)$$

4.4. Forward kinematics

The forward kinematics [27] deals with location and pose of the end effector with the joint variables and parameters. Here the end-effector frame is mapped to the base joint of the arm using homogenous reference frames as discussed above. Denavit–Hartenberg parameters [14], [11] (also called DH parameters) are the four parameters associated with a particular convention for attaching reference frames to the links of a spatial kinematic chain, or robot manipulator. These are calculated for all the joints using the Figure 10 as a reference and data given in Table 2. Origin O(i) is the intersection between x(i) and z(i) axes. The joint angle (θ(i)) is t, the angle from axis x(i-1) to x(i) about z(i) using the right-hand rule. In joint 3 it has an offset of 180-degree constant between x (2) and x(3) and in joint 4 it has an offset of 90-degree constant between x(3) and x(4). Link twist (α(i-1)) is the angle from z(i-1) to z(i) measured about x(i-1) using right hand rule. In Table2 link length (a(i-1)) is equal to the difference between z(i-1) and z(i) along the x(i) direction. Link offset (d(i)) in Table 2 is the difference between x(i-1) and x(i) along the z(i) axis. Using equation 8, which is the general representation of the transform matrix from one frame to another. So by using the data available in Table 2 transform matrix for link 0 and link 1 is given in T^1_0 in (8) Similarly, (9) is the transform for link 1 and link 2. $T^2_1, T^3_2, T^4_3, T^5_4, T^6_5, T^7_6$ These are also similar kinds of transform matrices with their respective links. Finally, the transform matrix from base joint to the end effector is derived by T^7_0 shown in (16). The orientation difference between the definitions of the gripper in URDF to the DH conventions so the rotation need to be done around y axis (17) and then need to be done around z axis (18) that is R_c shown in (19).

Table 2: DH-Parameters for the arm.

Links	O(i)	θ(i)	α(i-1)	a(i-1)	d(i)
0-1	1	θ1	0	0	0.15
1-2	2	θ2	-90	0	0.14
2-3	3	θ3+180	180	0.45	0.45
3-4	4	θ4+90	-90	0.04	0
4-5	5	θ5	-90	0	0.091
5-6	6	θ6	90	0.03	0.106
6-7	7	θ7	0	0	0.095

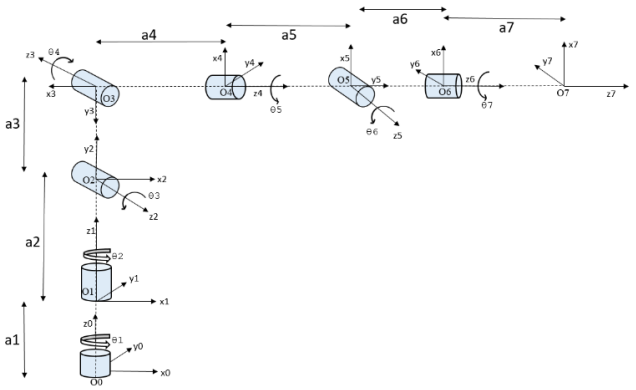


Figure 10: Cartesian representation of joints with rotational angles for DH-parameters.

$$T_i^{i-1}=\begin{bmatrix} c\theta_i & -s\theta_i & 0 & a_{i-1} \\ s\theta_i c\alpha_{i-1} & c\theta_i c\alpha_{i-1} & -s\alpha_{i-1} & -s\alpha_{i-1}d_i \\ s\theta_i s\alpha_{i-1} & c\theta_i s\alpha_{i-1} & c\alpha_{i-1} & c\alpha_{i-1}d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (8)$$

$$c\theta_i = \cos\theta_i$$

$$s\theta_i = \sin\theta_i$$

$$c\alpha_{i-1} = \cos\alpha_{i-1}$$

$$T_0^1=\begin{bmatrix} c\theta_1 & -s\theta_1 & 0 & 0 \\ s\theta_1 & c\theta_1 & 0 & 0 \\ 0 & 0 & 1 & 0\cdot 15 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

$$T_1^2=\begin{bmatrix} c\theta_2 & -s\theta_2 & 0 & 0 \\ 0 & 0 & 1 & 0\cdot 14 \\ -s\theta_2 & -c\theta_2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (10)$$

$$T_2^3=\begin{bmatrix} -c\theta_3 & s\theta_3 & 0 & 0.45 \\ s\theta_1 & c\theta_3 & 0 & 0 \\ 0 & 0 & -1 & -0\cdot 45 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (11)$$

$$T_3^4=\begin{bmatrix} s\theta_4 c\theta_4 & 0 & 0 & 0.04 \\ 0 & 0 & 1 & 0 \\ -c\theta_4 & -s\theta_4 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (12)$$

$$T_4^5=\begin{bmatrix} c\theta_5 & -s\theta_5 & 0 & 0 \\ 0 & 0 & 1 & 0.0911 \\ -s\theta_5 & -c\theta_5 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (13)$$

$$T_5^6=\begin{bmatrix} c\theta_6 & -s\theta_6 & 0 & 0.03 \\ 0 & c\theta_6 & -1 & -0.106 \\ c\theta_6 & c\theta_6 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (14)$$

$$T_6^7=\begin{bmatrix} c\theta_7 & -s\theta_7 & 0 & 0 \\ c\theta_7 & c\theta_7 & 0 & 0 \\ 0 & c\theta_7 & 1 & 0.095 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (15)$$

$$T_7^0 = T_1^0 * T_2^1 * T_3^2 * T_4^3 * T_5^4 * T_6^5 * T_7^6 \quad (16)$$

$$R_y=\begin{bmatrix} COS(-\pi/2) & 0 & SIN(\pi/2) & 0 \\ 0 & 1 & 0 & 0 \\ -SIN(-\pi/2) & 0 & COS(\pi/2) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (17)$$

$$R_z=\begin{bmatrix} COS(\pi) & -SIN(\pi) & 0 & 0 \\ SIN(\pi) & COS(\pi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (18)$$

$$R_c = R_y * R_z \quad (19)$$

4.5. Inverse kinematics (IK)

Inverse kinematics analysis is for obtaining the Joint angles by using end effector Cartesian space or position coordinates. Since the last three joints J4, J5, J6 in Figure5 and Figure 6 are revolute joints with joint axis intersection at J5 which would be wrist center (WC). Thus the kinematic of the IK is now evaluated by calculating Inverse position and Inverse Orientation.

4.5.1. Inverse Position

The inverse position problem is for obtaining the first three joint angles ($\theta_1, \theta_2, \theta_3$). To evaluate Inverse Position the end effector position (Px, Py, Pz) and Orientation (Roll, Pitch, Yaw) need to be taken from the test case input data. Thus the rotation matrix for the end effector R_rpy shown (20). After the error correction the actual end effector rotation matrix R_end is given in (21). In (22) the obtained matrix R_0^6 is the rotation part of the full homogeneous transform matrixT_0^{EE} which is a transform matrix of end effector to the base joint. Using these concepts the joint angles ($\theta_1, \theta_2, \theta_3$) are obtained.

$$R_rpy = Rot(Z, yaw) * Rot(Y, pitch) * Rot(X, Roll) \quad (20)$$

$$R_end = R_rpy * R_c \quad (21)$$

$$T_0^{EE}=\begin{bmatrix} r_{11} & r_{12} & r_{13} & p_x \\ r_{21} & r_{22} & r_{23} & p_y \\ r_{31} & r_{32} & r_{33} & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} R_0^6 & & & r^0{}_{0/EE} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (22)$$

4.5.2. Inverse Orientation

Inverse orientation problem is for obtaining the final three joint angles ($\theta_4, \theta_5, \theta_6$). The resultant transform is obtained using the individual DH transforms. Hence the resultant rotation is calculated as shown in (23). Since the overall the RPY (roll, pitch and yaw) rotation between base_link and gripper_link is equal to the product of individual rotations between the links as shown in (24). Multiply inverse matrix of R_3^0 (inv(R_3^0)) on either sides in (24) the resultant is shown in (25). The resultant matrix (25) on RHS (right hand side of the equation) does not have any variables after substituting the joint angles, and hence comparing LHS (left hand side of the equation) with RHS will result the equations for the last three joint angles ($\theta_4, \theta_5, \theta_6$). Solving the obtained equations which results the joint angles.

$$R_6^0 = R_1^0 * R_2^1 * R_3^2 * R_4^3 * R_5^4 * R_6^5 \quad (23)$$

$$R_end = R_6^0 \quad (24)$$

$$R_6^3 = \text{inv}(R_3^0)*\text{R_end} \quad (25)$$

5. Experiment and results

The robotic arm in the research is designed for the manipulation tasks for various applications such as Coconut crop harvesting, Assistive robots like self-driving wheel chairs and Home robots, Search and Rescue robots for disastrous environments, Collaborative robots for research use and etc. The evaluation of the design and software was first done on RVIZ simulation with few goal positions, these goal positions are the most desired positions of the end effector during the general operations. Once the software simulation has given satisfactory results, the testing was done on hardware which was already built. The designed 6 DOF robotic arm is simulated in RVIZ. An axis plugin has been added to visualize a coordinate axis at the base of the model robotic arm. The blue line indicates the z-axis, the red line indicates the x-axis and green line indicates the y-axis. The silver-coloured arm represents the goal pose of the arm for the user input values.

Figure 11 represents the different poses of the robotic arm when different coordinates are input to IK solver. Different x, y, and z coordinates are published to the interface module to manipulate the pose of the model arm. Table 3, 5 and 6 represents the outcomes of the IK solver when different goal points are input into the solver using different control methods on the RVIZ simulation platform. The output joint state values of all the goals positions in Figure 11 are given in Table 3, 5 and 6. The input point is considered with respect to the RVIZ [17] world axis frame. With the IK based approach, the robotic arm was manipulated at the desired points in the RVIZ world frame simulation. The three control modes have been tested on 16 different goal positions. The input data from the interface block enters the IK solver node where the required joint values are calculated to achieve the respective input pose. These generated joint values are passed to Joint state publisher and robot state publisher where the required Tf and combined joint state array will be passed to the RVIZ for the simulation.

On comparing the accuracy between the three methods keyboard and GUI were able to position the arm to the desired goal pose. Joystick failed in some cases as the control mode was difficult to manipulate the robotic arm. Since the Joystick PWM values need to be transferred to ROS via a microcontroller, due to the disturbances in the hardware, in a few cases joystick input readings varied. There was 1%-3% error in the input and hence the output was also affected due to the change in the input values. Comparing the results from Table 3, 4, 5 it is clear that GUI and Keyboard are two ideal methods to accurately position the robotic arm at a desired position.

The Figure 12 and Figure 13 illustrates the error percentage values while conducting a simulation experiment through a joystick. A survey has been conducted by testing the system with a 20 number of individuals. Based on the outputs that were obtained, an analysis has been made on the error percentage in reaching the goal position. The error is calculated by comparing the correct outcomes obtained by an individual to the total no of test cases in Figure 12. And an analysis has been made which showed the error percentage for each goal pose. This is calculated based on the pass and fail cases obtained while conducting the survey by 20 individuals. The test results are plotted in Figure 13. Considering these errors, joystick control was avoided for the hardware test. The hardware test has been implemented using GUI as a control interface since the simulation results of keyboard based control and GUI based control are the same. Table 6 represents the outcomes of IK solver when implemented on hardware using GUI based control. It is observed that the output in the hardware testing using GUI based control is similar to the simulation testing using GUI.

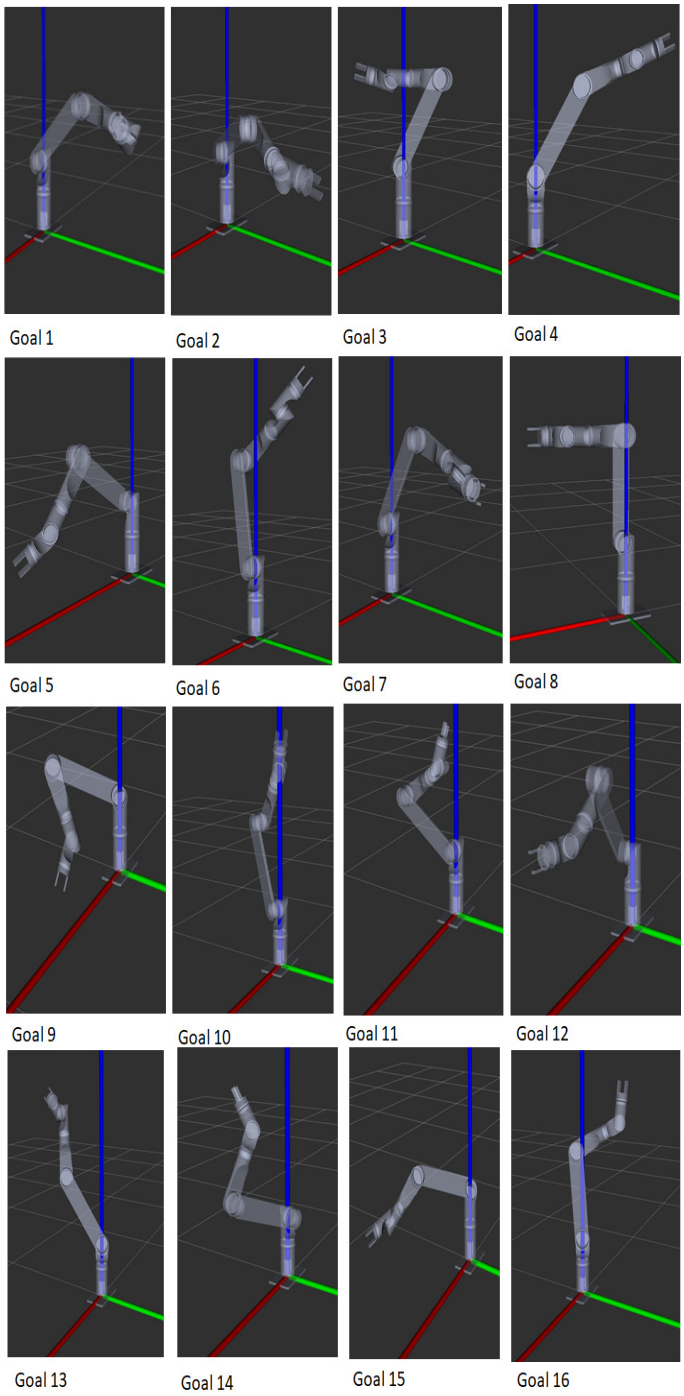


Figure 11: The goal pose of the robotic arm on different user inputs.

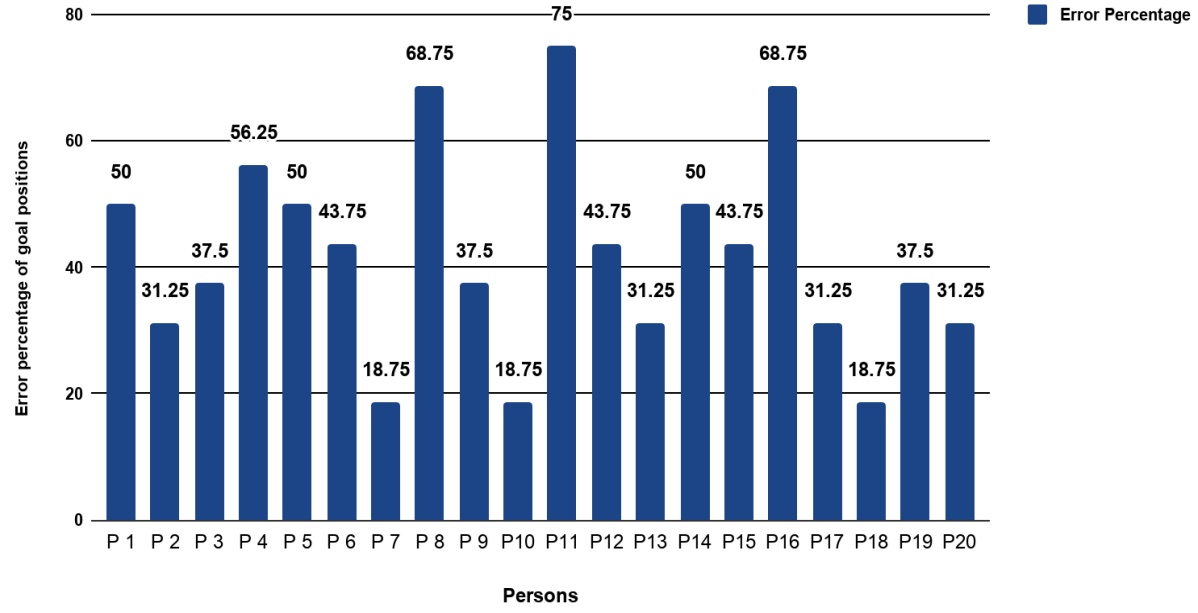


Figure 12: Error percentage for individual person

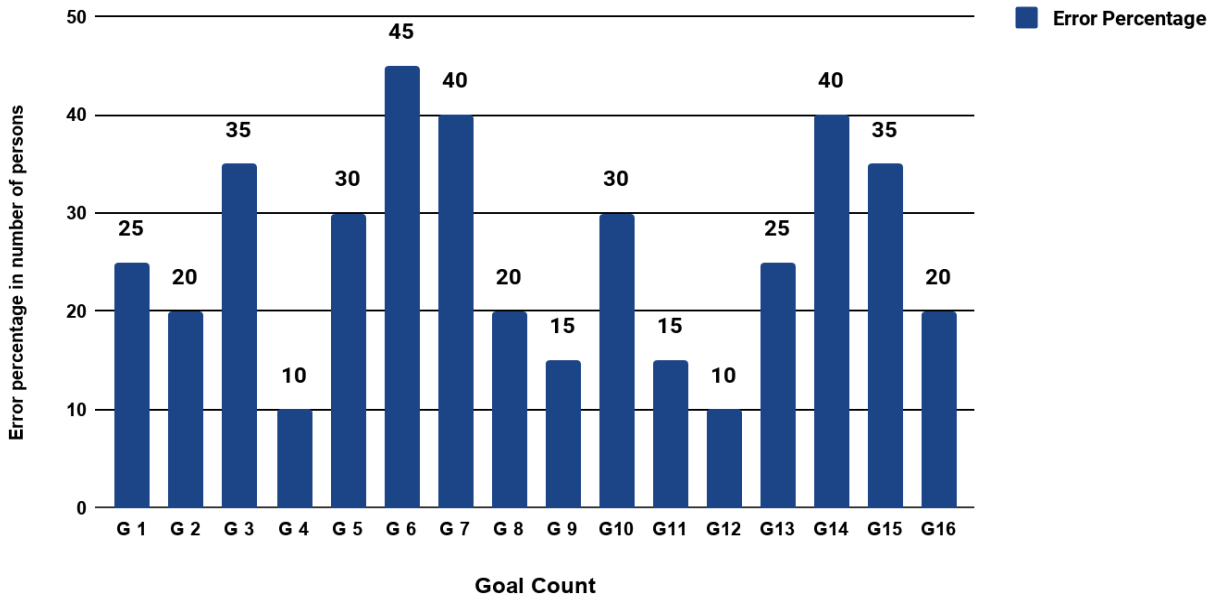


Figure 13 Error Percentage for individual goal position

Table 3: Outcomes of the joint state values from the IK solver node in simulation for the user input coordinates using GUI based control.

Goal Count	X position	Y position	Z position	DOF1 (radians)	DOF2 (radians)	DOF3 (radians)	DOF4 (radians)	DOF5 (radians)	DOF6 (radians)
Goal 1	0.301	0.651	0.525	1.2704	-0.9539	0.7554	-1.0660	1.3424	-1.185
Goal 2	0.402	0.730	0.376	1.0098	-1.222	0.9088	1.2786	0.476	1.4327
Goal 3	0.133	- 0.116	0.699	-0.5268	0.5868	-0.567	-0.6876	0.3067	0.0028
Goal 4	-0.392	0.532	0.825	2.2363	-0.6110	0.9545	0.043966	0.0151	0.6119
Goal 5	0.675	- 0.095	0.220	-0.0797	-1.0676	0.2312	-0.8217	0.4022	0.9090
Goal 6	- 0.019	0.242	1.020	1.4644	0.1755	0.3900	1.0424	0.5786	-1.297
Goal 7	0.204	0.644	0.529	1.3112	0.6556	0.3388	-0.9354	0.2892	1.14233
Goal 8	0.428	- 0.020	0.760	0	0	0	0	0	0

Goal 9	0.255	- 0.267	0.067	-0.89364	-1.23433	-0.70224	0.579329	0.924559	-0.41793
Goal 10	0.043	0.027	1.079	1.12537	0.424841	0.580440	-0.63930	0.2128	-0.99506
Goal11	- 0.017	- 0.110	0.759	1.396672	1.164312	-0.62972	0.222626	0.77854	0.505854
Goal 12	0.671	- 0.020	0.510	0.028259	-0.67949	0.223439	-0.42861	0.52486	0.095769
Goal 13	0.141	- 0.338	1.038	2.177904	0.639303	0.979159	1.05975	1.08808	-0.56990
Goal 14	0.281	- 0.123	0.844	2.831651	1.213924	-0.16744	0.57619	0.93674	-1.19414
Goal 15	0.303	- 0.716	0.108	-1.25223	-1.48647	0.71624	1.17593	0.725	-0.75893
Goal 16	-0.116	0.235	1.001	2.050420	0.0916879	0.3374000	0.156057999	1.03124	0.649038000

Table 4: Outcomes of the joint state values from the IK solver node in simulation for the user input coordinates using Keyboard.

Goal Count	X position	Y position	Z position	DOF1 (radians)	DOF2 (radians)	DOF3 (radians)	DOF4 (radians)	DOF5 (radians)	DOF6 (radians)
Goal 1	0.301	0.651	0.525	1.2704	-0.9539	0.7554	-1.0660	1.3424	-1.185
Goal 2	0.402	0.730	0.376	1.0098	-1.222	0.9088	1.2786	0.476	1.4327
Goal 3	0.133	- 0.116	0.699	-0.5268	0.5868	-0.567	-0.6876	0.3067	0.0028
Goal 4	-0.392	0.532	0.825	2.2363	-0.6110	0.9545	0.043966	0.0151	0.6119
Goal 5	0.675	- 0.095	0.220	-0.0797	-1.0676	0.2312	-0.8217	0.4022	0.9090
Goal 6	- 0.019	0.242	1.020	1.4644	0.1755	0.3900	1.0424	0.5786	-1.297
Goal 7	0.204	0.644	0.529	1.3112	0.6556	0.3388	-0.9354	0.2892	1.14233
Goal 8	0.428	- 0.020	0.760	0	0	0	0	0	0
Goal 9	0.255	- 0.267	0.067	-0.89364	-1.23433	-0.70224	0.579329	0.924559	-0.41793
Goal 10	0.043	0.027	1.079	1.12537	0.424841	0.580440	-0.63930	0.2128	-0.99506
Goal 11	- 0.017	- 0.110	0.759	1.396672	1.164312	-0.62972	0.222626	0.77854	0.505854
Goal 12	0.671	- 0.020	0.510	0.028259	-0.67949	0.223439	-0.42861	0.52486	0.095769
Goal 13	0.141	- 0.338	1.038	2.177904	0.639303	0.979159	1.05975	1.08808	-0.56990
Goal14	0.281	- 0.123	0.844	2.831651	1.213924	-0.16744	0.57619	0.93674	-1.19414
Goal 15	0.303	- 0.716	0.108	-1.25223	-1.48647	0.71624	1.17593	0.725	-0.75893
Goal16	-0.116	0.235	1.001	2.050420	0.0916879	0.3374000	0.156057999	1.03124	0.649038000

Table 5: Outcomes of the joint state values from the IK solver node in simulation for the user input coordinates using Joystick.

Goal Count	X position	Y position	Z position	DOF1 (radians)	DOF2 (radians)	DOF3 (radians)	DOF4 (radians)	DOF5 (radians)	DOF6 (radians)
Goal 1	0.305	0.652	0.525	1.2704	-0.9540	0.7556	-1.0660	1.3424	-1.185
Goal 2	0.406	0.7304	0.376	1.0098	-1.230	0.9089	1.2786	0.476	1.4327
Goal 3	0.135	- 0.116	0.699	-0.5268	0.5871	-0.567	-0.6876	0.3067	0.0028
Goal 4	-0.394	0.530	0.825	2.2363	-0.6116	0.9548	0.043966	0.0151	0.6119
Goal 5	0.67	- 0.095	0.220	-0.0797	-1.0681	0.2312	-0.8217	0.4022	0.9090
Goal 6	- 0.015	0.242	1.020	1.4644	0.1755	0.3900	1.0424	0.5786	-1.297
Goal 7	0.209	0.645	0.529	1.3112	0.6558	0.3389	-0.9354	0.2892	1.14233
Goal 8	0.421	- 0.020	0.760	0	0	0	0	0	0
Goal 9	0.255	- 0.267	0.067	-0.89364	-1.23433	-0.70224	0.579329	0.924559	-0.41793
Goal 10	0.043	0.028	1.079	1.12537	0.424841	0.58048	-0.63930	0.2128	-0.99506
Goal11	- 0.017	- 0.110	0.759	1.396672	1.164312	-0.62972	0.222626	0.77854	0.505854
Goal 12	0.671	- 0.020	0.510	0.028259	-0.67949	0.223439	-0.42861	0.52486	0.095769
Goal 13	0.141	- 0.338	1.038	2.177904	0.639303	0.979159	1.05975	1.08808	-0.56990
Goal14	0.281	- 0.123	0.844	2.831651	1.213924	-0.16744	0.57619	0.93674	-1.19414
Goal 15	0.303	- 0.716	0.108	-1.25223	-1.48647	0.71624	1.17593	0.725	-0.75893
Goal16	-0.116	0.234	1.001	2.050420	0.0916879	0.3375	0.156057999	1.03124	0.649038000

Table 6: Outcomes of the joint state values from the IK solver node on hardware testing for the user input coordinates using GUI

Goal Count	X position	Y position	Z position	DOF1 (radians)	DOF2 (radians)	DOF3 (radians)	DOF4 (radians)	DOF5 (radians)	DOF6 (radians)
Goal 1	0.301	0.651	0.525	1.2799	-0.9538	0.7600	-1.0659	1.3426	-1.184
Goal 2	0.402	0.730	0.376	1.0095	-1.222	0.9086	1.2786	0.480	1.4327
Goal 3	0.133	- 0.116	0.699	-0.5268	0.5868	-0.567	-0.6876	0.3067	0.0028
Goal 4	-0.392	0.532	0.825	2.2363	-0.611015	0.95467	0.0439666	0.01561	0.6119
Goal 5	0.675	- 0.095	0.220	-0.07978	-1.0676	0.2312	-0.82172	0.4022	0.9090
Goal 6	- 0.019	0.242	1.020	1.464454	0.1755	0.3900	1.0424	0.57862	-1.29711
Goal 7	0.204	0.644	0.529	1.3112	0.6556	0.3388	-0.9354	0.2892	1.14233
Goal 8	0.428	- 0.020	0.760	0	0.	0	0	0	0
Goal 9	0.255	- 0.267	0.067	-0.89364	-1.234335	-0.702246	0.579329	0.924559	-0.41793
Goal 10	0.043	0.027	1.079	1.12537	0.424841	0.5804405	-0.63930	0.21285	-0.99506
Goal11	- 0.017	- 0.110	0.759	1.3966725	1.1643121	-0.629728	0.2226262	0.77854	0.505854
Goal 12	0.671	- 0.020	0.510	0.028259	-0.67949	0.223439	-0.42861	0.52486	0.095769
Goal 13	0.141	- 0.338	1.038	2.177904	0.639303	0.97915954	1.0597566	1.088084	-0.56990
Goal14	0.281	- 0.123	0.844	2.831651	1.2139244	-0.16744	0.576198	0.93674	-1.19414584
Goal 15	0.303	- 0.716	0.108	-1.252234	-1.48647	0.7162445	1.175932	0.72512	-0.75893
Goal 16	-0.116	0.235	1.001	2.050420	0.0916879	0.3374000	0.156057999	1.03124	0.649038000

6. Conclusion

In this research work, the authors proposed and evaluated reliable methods for controlling a robotic arm by testing it in both hardware and simulation using RVIZ. The authors in [17] used the MoveIt to build a kinematics library for the IK of the robotic arm, but as an extension of that, this research designed and developed a kinematics-based solution using DH parameters method. Using the derived equations, successfully tested the designed robotic arm in Rviz using the proposed control mechanisms. A survey is conducted in evaluating the best control methodology where Figure [12] and Figure [13] depict the results. Based on the survey, the research validates that the joystick failed in achieving the desired input coordinate positions due to signaling and hardware issues. But GUI and Keyboard showed better results in controlling the arm.

As a rule, simulations do not reproduce the exact real-time behavior of an entity or a system. PID based control can reduce the error. The research can be enhanced by testing the proposed design in the Gazebo simulation software. Gazebo provides the flexibility to use a PID-based controller, which helps in smooth and exact mimicking of the real-time robotic arm as per simulation in RVIZ. The proposed testing gives the developer good results. Design of the end effector can be improved for performance of multiple, divergent tasks in a real time environment. The singularity issues can be reduced for better performance and enhancement of the task.

Acknowledgment

We are grateful for Amrita Vishwa Vidyapeetham and Humanitarian and Technology Labs for providing us with all the sophisticated requirements to develop and complete this paper.

References

[1] Megalingam R.K, Sivanantham V, Kumar K.S, Ghanta S, Teja P.S, Gangireddy R, Sakti Prasad K.M, Gedela V.V, "Design and development of inverse kinematic based 6 DOF robotic arm using ROS", International Journal of Pure and Applied Mathematics, 118, pp. 2597-2603, ISSN: 1314-3395

[2] Xinquan Liang, Haris Cheong, Yi Sun, Jin Guo, Chee Kong Chui, Chen-Hua Yeo, "Design, Characterization, and Implementation of a Two-DOF Fabric-Based Soft Robotic Arm", IEEE Robotics and Automation letters(Volume:3,Issue:3,July2018),https://doi.org/10.1109/LRA.2018.2831723

[3] Asghar Khan, Wang Li Quan. "Structure design and workspace calculation of 6-DOF underwater manipulator". 2017 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST), 10-14 Jan 2014, https://doi.org/ 10.1109/IBCAST.2017.7868119

[4] Elkin Yesid Veslin; Max Suell Dutra; Omar Lengerke; Edith Alejandra Carreño; Magda Judith Morales, "A Hybrid Solution for the Inverse Kinematic on a Seven DOF Robotic Manipulator", IEEE Latin America Transactions (Volume: 12, Issue: 2, March 2014), https://doi.org/ 10.1109/TLA.2014.6749540

[5] Pietro Falco; Ciro Natal, "On the Stability of Closed-Loop Inverse Kinematics Algorithms for Redundant Robots", 05 May 2011, IEEE Transactions on Robotics (Volume: 27, Issue: 4, Aug. 2011),https://doi.org/ 10.1109/TRO.2011.2135210

[6] Seo-Wook Park, Jun-Ho Oh, "Hardware Realization of Inverse Kinematics for Robot Manipulators", IEEE Transactions on Industrial Electronics Vol.41, No 1, February 1994, https://doi.org/ 10.1109/41.281607

[7] Rajesh Kannan Megalingam, Nigam Katta, Raviteja Geesala, Prasant Kumar Yadav, "Keyboard Based Control and Simulation of 6 DOF Robotic Arm using ROS", 2018 4th International Conference on Computing Communication and Automation(ICCCA), https://doi.org/10.1109/CCAA.2018.8777568

[8] Wei Qian, Ziyang Xia, Jing Xiong, Yangzhou Gan, Yangchow Guo, Shaokui Weng, Hao Deng, Ying Hu, Jiawei Zhang, "Manipulation Task simulation using ROS and Gazebo", Robotics and Biomimetics (ROBIO), 2014 IEEE International Conference", https://doi.org/10.1109/ROBIO.2014.7090732

[9] Žlajpah Leon, Simulation in robotics (2008) Mathematics and Computers in Simulation, 79 (4), 15 December 2008, pp. 879-897, https://doi.org/10.1016/j.matcom.2008.02.017

[10] Kichang Lee ; Jiyoung Lee ; Bungchul Woo ; Jeongwook Lee ; Young-Jin Lee, Kimhae-si ; Syungkwon Ra "Modeling and Control of an Articulated Robot Arm with Embedded Joint Actuators ",2018 International Conference on Information and Communication Technology Robotics (ICT-ROBOT), https://doi.org/10.1109/ICT-ROBOT.2018.8549903

[11] Alla N Barakat, Khaled A. Gouda, Kenza.Bozed, " Kinematics analysis and simulation of a robotic arm using MATLAB",2016 4th International Conference on Control Engineering & Information Technology (CEIT),16-18 Dec. 2016, Hammamet, Tunisia, https://doi.org/ 10.1109/CEIT.2016.7929032

[12] Stefano Carpin, Mike Lewis, Jijun Wang, "USARSim: a robot simulator for research and education" Proceedings 2007 IEEE International Conference on Robotics and Automation, 10-14 April 2007, Roma Italy, https://doi.org/ 10.1109/ROBOT.2007.363180

[13]] Jun-Di Sun; Guang-Zhong Cao; Wen-Bo Li; Yu-Xin Liang; Su-Dan Huang, "Analytical inverse kinematic solution using the D-H method for a 6-DOF robot", 2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), https://doi.org/10.1109/URAI.2017.7992807

[14] Akhilesh Kumar Mishra; Oscar Meruvia-Pastor; " Robot arm manipulation using depth-sensing cameras and inverse kinematics "2014 Oceans-St John's,14-19 September 2014 IEEE International Conference, St. John's, NL, Canada, https://doi.org/ 10.1109/OCEANS.2014.7003029

[15] C. A. G. Gutiérrez; J. R. Reséndiz; J. D. M. Santibáñez; G. M. Bobadilla, "A Model and Simulation of a Five-Degree-of-Freedom Robotic Arm for Mechatronic Courses", IEEE Latin America Transactions (Volume: 12, Issue: 2, March 2014) ,https://doi.org/ 10.1109/TLA.2014.6749521

[16] Zenghua Bian ; Zhengmao Ye ; Weilei Mu," Kinematic analysis and simulation of 6-DOF industrial robot capable of picking up die-casting products", 2016 IEEE International Conference on Aircraft Utility Systems (AUS), https://doi.org/ 10.1109/AUS.2016.7748017

[17] S. Hernandez-Mendez, C. Maldonado-Mendez, A. Marin-Hernandez, H. V. Rios-Figureueroa, H. Vazquez-Leal and E. R. Palacios-Hernandez, "Design and implementation of a robotic arm using ROS and MoveIt!", 2017 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), Ixtapa, 2017, pp. 1-6, http://doi.org/ 10.1109/ROPEC.2017.8261666

[18] Weimin Shen, Jason Gu, Yide Ma, "3D Kinematic Simulation for PA10-7C Robot Arm Based on VRML", 18-21 Aug. 2007, Jinan, China. https://doi.org/10.1109/ICAL.2007.4338637

[19] José de Jesús Rubio, "Sliding mode control of robotic arms with dead zone", IET Control Theory & Applications (Volume: 11, Issue: 8, 5 12 2017), https://doi.org/ 10.1049/iet-cta.2016.0306

[20] Megalingam R.K, Rajesh Gangireddy, Gone Sriteja, Ashwin Kashyap, Apuroop Sai Ganesh "Adding intelligence to the robotic coconut tree climber" 2017 International Conference on Inventive Computing and Informatics (ICICI), 23-24 Nov. 2017, Coimbatore India, https://doi.org/ 10.1109/ICICI.2017.8365206

[21] Teerawat Thepmanee, Jettiya Sripituk, Prapart Ukakimapurn, "A simple technique to modeling and simulation four-axe robot-arm control" 17-20 Oct. 2007, Coex, Seoul, Korea, https://doi.org/10.1109/ICCAS.2007.4406694

[22] Tingting Meng; Wei He, "Iterative Learning Control of a Robotic Arm Experiment Platform with Input Constraint", IEEE Transactions on Industrial Electronics (Volume: 65, Issue: 1, Jan. 2018), https://doi.org/ 10.1109/TIE.2017.2719598

[23] Gourab Sen Gupta, Subhas Chandra Mukhopadhyay, Christopher H. Messom, Serge N. Demidenko, "Master-Slave Control of a Teleoperated Anthropomorphic Robotic Arm With Gripping Force Sensing ", IEEE Transactions on Instrumentation and Measurement (Volume: 55, Issue: 6, Dec. 2006),https://doi.org/ 10.1109/TIM.2006.884393

[24] P. Lucibello," Repositioning control of robotic arms by learning ", IEEE Transactions on Automatic Control (Volume: 39, Issue: 8, Aug 1994), https://doi.org/ 10.1109/9.310053

[25] Qingyuan Sun; Lingcheng Kong; Zhihua Zhang; Tao Mei, "Design of wireless sensor network node monitoring interface based on Qt", 2010 International Conference on Future Information Technology and Management Engineering, Changzhou, 2010, pp. 127-130

[26] R. Takuma; Y. Asahara ; H. Kajimoto ; N. Kawakami ; S. Tachi, "Development of anthropomorphic multi-D.O.F master-slave arm for mutual telexistence ", IEEE Transactions on Visualization and Computer Graphics (Volume: 11 , Issue: 6 , Nov.-Dec. 2005), https://doi.org/ 10.1109/TVCG.2005.99

[27] Ramesh, S. B. Hussain and F. Kangal, "Design of a 3 DOF robotic arm," 2016 Sixth International Conference on Innovative Computing Technology (INTECH), Dublin, 2016, pp. 145-149, https://doi.org/ 10.1109/INTECH.2016.7845007

[28] Jose de Jesus Rubio; Adrian Gustavo Bravo; Jaime Pacheco; Carlos Aguilar,"Passivity analysis and modeling of robotic arms", IEEE Latin America Transactions (Volume: 12 , Issue: 8 , Dec. 2014) ,https://doi.org/ 10.1109/TLA.2014.7014505

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 307-316 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Effects of Split Ring Resonator (SRR) Metamaterial on the Radiation Pattern and Variation of the Heating Focus Point of the Microstrip Patch Antenna

Mahsa Jahangiri*, Saman Rajebi

Electrical Engineering Department, Seraj Higher Education Institute, Tabriz, Iran

ARTICLE INFO

Article history:
Received: 03 September, 2019
Accepted: 13 January, 2020
Online: 07 February, 2020

Keywords:
Microstrip patch antenna
Metamaterial
Split ring resonator
Circular hole
Hyperthermia
Specific absorption rate (SAR)

ABSTRACT

Metamaterial antennas are some sort of antennas which utilize metamaterials to improve the efficiency of small (small electrical) antenna systems and also enhance the antenna's radiation pattern. The unusual properties of metamaterials such as coefficients of permittivity and permeability, as well as the negative refractive index, have led to increase the use of these artificial materials in variant fields including microwave engineering and medical applications. This paper explores a metamaterial coating and impact of it on a microstrip patch antenna from two different perspectives on satellite and biomedical applications. For this purpose, a microstrip antenna is modeled using CST software and then a particular metamaterial cover is configured upon the patch antenna to enhance its efficiency such as radiation pattern and directivity. The method is arranged in several stages and the output of the antenna is then analyzed in each stage after applying some adjustments to the design process. On the other hand, because the type and position of the applicator on the tissue is a consequential variable, in order to improve the effectiveness of hyperthermia treatment in medical applications, in the next step the metamaterial antenna is applied to hyperthermia treatment on the body tissue. Eventually the results of metamaterial antenna are compared with the usual antenna.

1. Introduction

Current paper is an extended version of work originally presented in [1].

Veselago first proposed the concept of the negative refractive index and the presence of substances with this feature in 1968 [2]. By using a wired structure John Pandari succeeded in achieving negative permittivity in 1996, and by using split resonator rings, he then prospered independently in generating an ambit with negative permeability in 1999 [3]. The idea was fulfilled experimentally and actually in 2000 by David Smith and his fellow members, applying an intermittent volumetric series of alternative metal wires and split ring resonator combinations [4,5]. The name given to these materials is "Metamaterial".

Since the metamaterials have the negative ε and μ, n has a negative root in certain frequencies which is formulated by Maxwell equation that is $n = \pm\sqrt{\varepsilon\mu}$ [6]. The peculiar and appealing features of these materials have contributed to their usage in diverse fields, such as microwave engineering, waveguide operation, phase compensation, dispersion correction, smart new antennas, lenses and numerous other instances. A type of metamaterial structure and its impacts on performance of a microstrip antenna will be studied in this article.

If an electrical field is applied in parallel to a structure composed of a series of thin metal rods, a negative permittivity property will be revealed in the structure. Likewise, if a magnetic field is perpendicularly exerted to a mechanism with a set of split ring resonators (SRR), it will exhibit a negative permeability [7].

The Microstrip Antenna Concept was first presented in 1950 [8]. The reasons for its success were proper characteristics including lower volume, light weight, low expense and reasonable efficiency [9,10]. A plan is proposed in this paper that using it we will explore the capacities as well as restrictions of metamaterial mechanism in antenna design.

The standard microstrip antenna is formed of three segments of ground, a radiation surface and, a dielectric substrate. This

*Corresponding Author: Mahsa Jahangiri, Seraj Higher Education Institute, Tabriz, Iran, Email: jahangiri.mahsa66@gmail.com

https://dx.doi.org/10.25046/aj050139

dielectric substratum lies between the patch (radiation surface) and the ground layer.

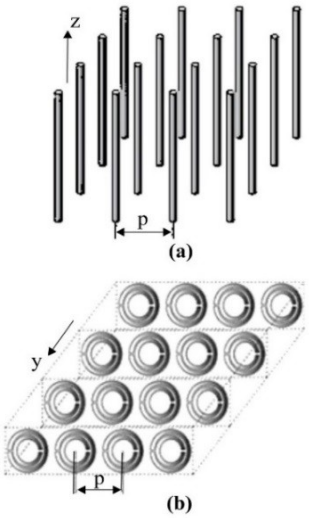


Figure 1: An illustration of intermittent patterns. (a) The metal thin rods, (b) The split ring resonators (SRRs)

Equations (1), (2), (3), (4) and (5) measure the sizes of a standard patch, where W and L are width and length of the radiation patch, respectively [9,11].

$$\varepsilon_{reff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r + 1}{2}\left[1 + 12\frac{h}{W}\right]^{-\frac{1}{2}} \quad (1)$$

$$W = \frac{c}{2f_0\sqrt{\frac{(\varepsilon_r + 1)}{2}}} \quad (2)$$

$$L_{eff} = \frac{c}{2f_0\sqrt{\varepsilon_{reff}}} \quad (3)$$

$$\Delta L = 0.412h\frac{(\varepsilon_{reff} + 0.3)(\frac{W}{h} + 0.264)}{(\varepsilon_{reff} - 0.258)(\frac{W}{h} + 0.8)} \quad (4)$$

$$L = L_{eff} - 2\Delta L \quad (5)$$

The theory of the effective quantity has been used here. We need data of the scattering parameters (S_{11} and S_{21}) for a wave incident normally on a finite slab of a homogeneous material, to recover its electromagnetic characteristics including analytical quantities of permeability (μ) and permittivity (ε), that express a straightforward description. S_{11} and S_{21} are measured via (6) and (7) [12]:

$$S_{11} = \frac{i}{2}\left(\frac{1}{z} - z\right) sin(nkd)\, S_{21} \quad (6)$$

$$S_{21} = \frac{1}{cos(nkd) - \frac{i}{2}(z + \frac{1}{z})\, sin(nkd)} \quad (7)$$

Also the permeability and permittivity are computed by the relations μ = nz and ε = n / z [12, 13].

As it is evident, scattering parameters are correlated with ε and μ, and also with n and z. Hence, by inverting the S-parameters, the refractive index (n) and impedance (z) would be derived. For the calculation of n and z, equations (8) [13] and (9) are used [14]:

$$z = \sqrt{\frac{(1 + s_{11})^2 + {s_{21}}^2}{(1 - s_{11})^2 - {s_{21}}^2}} \quad (8)$$

$$n = \frac{1}{kd}cos^{-1}\left[\frac{1}{2s_{21}}(1 - {s_{11}}^2 + {s_{21}}^2)\right] \quad (9)$$

In which $k=\frac{w}{c}$ indicates the wave number of free space and d is the slab thickness [12].

1.1 Overview of hyperthermia

Cancer rates are increasing alarmingly all around the world, and this matter is a major cause of concern for the medical community [15]. Cancer tumors are able to affect different tissues and spread to different body organs. In recent years, in order to deal with the disease, researchers have developed a variety of treatments, including surgery, chemotherapy, radiotherapy, immunotherapy, stem cell transplant, and targeted therapies. Beside these methods, another well-known method called hyperthermia (or heat treatment) has recently been evaluated in clinical trials. Hyperthermia is a treatment method in which tumor tissues are exposed to high temperatures to be eliminated, without affecting healthy tissues [16]. Hyperthermia can be performed in both invasive and non-invasive methods. The non-invasive hyperthermia treatment is an efficacious safe treatment with less pain and injury to counteract different types of cancer [17]. One of the biggest challenges in the non-invasive treatment of the hyperthermia, is the concentration of electromagnetic energy on the cancerous tissue in a way that not to damage to healthy tissues. In this treatment, it is required that the temperature of the treated area be raised from 35 ° C to about 42 to 45 ° C [16]. During the period of hyperthermia treatment, tumor cells are exposed to a temperature of about 42 ° C for an hour or even longer [15]. Hence, many researchers have recently focused on non-invasive hyperthermia antennas and applicators. The applicator has been considered as the principal part of the hyperthermia which uses electromagnetic energy to be transmitted and distributed in the form of heat. It is then absorbed by cancer tissues to a certain level of heat and time [17].

1.1.1 SAR in hyperthermia

The quantity that is investigated in hyperthermia topic, is a parameter called specific absorption rate (SAR). The SAR quantity represents the quantity of heat energy absorbed by living tissue per unit time, when the rise of body temperature due to an external heat source exceeds one degree celsius. Specific absorption rate is defined as the energy absorbed per unit mass of tissue in watts per kilogram by (10):

$$SAR = \frac{d}{dt}\frac{dw}{dm} = \frac{d}{dt}\frac{dw}{\rho dv} \quad (10)$$

The SAR measurement can also be expressed in terms of the tissue temperature rise by (11):

$$SAR = C_h \frac{dT}{dt} \quad (11)$$

The SAR can also be calculated as (12):

$$SAR = \frac{\sigma E^2}{\rho} \quad (12)$$

which in the (10), (11) and (12), E is intensity of the electric field in the tissue with unit of volt/meter, σ is conductivity of tissue liquid with unit of Siemens per meter, ρ is volumetric density of tissue with unit of kg/m3, C_h is specific heat capacity of tissue with unit of J/kg and Kelvin, dT/dt is primary time derivative of tissue temperature in Kelvin unit. Therefore, SAR is calculated by measuring the amplitude of the electric field intensity [18].

2. Methodology

2.1. Original thought behind current study

This article's original concept was taken from [19]. A microstrip patch antenna was proposed in the aforesaid paper, the radiation pattern was then ameliorated with the help of a metamaterial coating composed of split ring resonators (SRR) and wires that generate negative permeability and permittivity. Table 1 represents the findings of the above-mentioned paper.

It is worth mentioning that the SRR (Split-Ring Resonator) unit cell structure was derived from [20]. The distinction between them is that the unit cell was regarded as a singular square in the mentioned article while it is constructed from a couple of squares adjacent to one another in the present study.

Table 1: Compared results in paper [19] between the usual patch antenna and the metamaterial antenna

Parameters	Usual patch antenna	Metamaterial antenna
Return losses (dB)	-30	-23
Gain(dB)	6.80	7.32
HPBW(3dB)	80°	76°

2.2. Stages of configuration of the metamaterial antenna

To obtain the desired frequency in the 8-12 GHz range, the patch measurements (L × W) are calculated as (11.86 mm × 9.06 mm) and for the ground sizes (Lg × Wg) are measured as (21.34 mm × 18.54 mm). The frequency of resonance for the proposed antenna is 9.6 GHz [1].

The coaxial cable has been considered as the feeding of the aimed Microstrip patch antenna. In order to properly match with the input impedance, coaxial feeding can be positioned within the patch at any point [10]. To determine the location of the coaxial cable precisely, the input impedance is adjusted to: $cos^2(\frac{\pi y_0}{L})$ by 50 ohms.

Figure 2 displays a general form of an ordinary rectangular patch antenna which is stimulated by coaxial line.

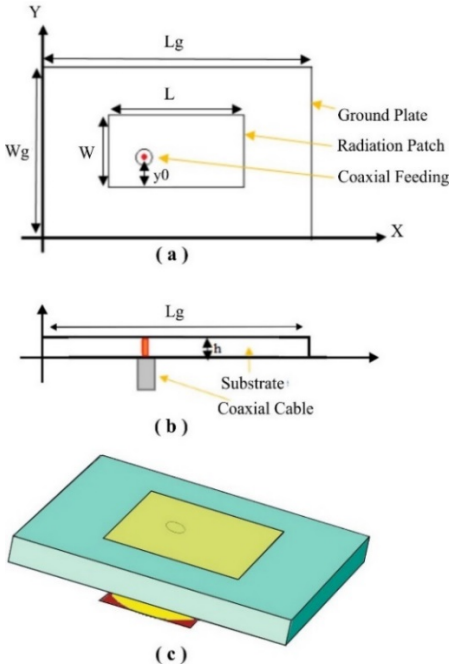


Figure 2: The microstrip patch antenna designed for 9.6 GHz. (a) Top-view geometry, (b) Side-view geometry, (c) Software-designed antenna

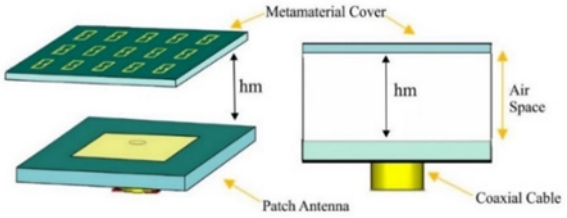


Figure 3: The metamaterial-coated microstrip patch antenna

The dielectric segment of the antenna has been constructed of Epsilon 2.2 substratum of Teflon and 1.58 mm thickness. Copper has been used to make the ground plane and radiation patch. In this paper, a cover of metamaterial has been utilized over a patch antenna as shown in Figure 3. This design forms of a 0.5 mm Teflon sheet of epsilon 2.2, which has 5*3 periodic slotted rings printed on the surface of this Teflon coating (single layer) measuring 0.017 mm thick and built of copper. This system consists of simple elements which are extremely expandable. The metamaterial layer is positioned from the patch of antenna at a distance of hm=10.53 mm, which is an air gap. This air distance was assumed to be zero at first, and the metamaterial was located directly on the antenna. Then in order to compare the changes, an air distance between the antenna and the metamaterial was regarded. The comparisons and findings have been provided in the section 3. In another phase, the SRR's intermittent configuration was arranged in two layers which were mounted on both sides of the Teflon coating surfaces in equal numbers. The findings have been analyzed in section 3 and compared to the singular layer SRR type. During other stage, we modified the unit cells and attained new results. At the end, the function of metamaterial antenna in hyperthermia treatment, has been investigated.

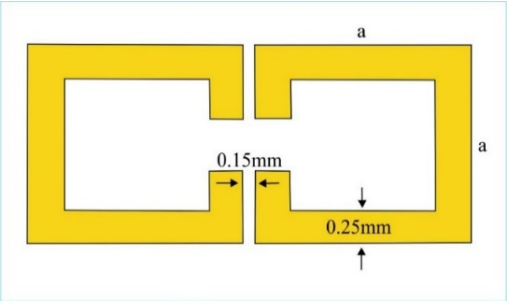


Figure 4: The designed SRR unit cell

Figure 3 displays the common microstrip antenna with the placement of metamaterial cover and Figure 4 illustrates the unit cell of SRR. As shown in Figure 4, there are two square-like components of the unit cell with a distance between them that its size is 0.15 mm. Line width of each component is 0.25 mm and the length is a=1.46 mm [1].

During the first phase, the outcomes of the antenna without metamaterial coating have been explored, taking into account the measurements and configuration of the coaxial feeding microstrip patch antenna.

In general, the real length of the patch in this group of microstrip antennas is the half of the wavelength, it has a weak directivity of almost 5 to 6 dB and a half-power bandwidth scope of around 70 to 90 degrees [21]. In this study, the purpose of applying metamaterial cover is to boost gain and beam direction. The aimed spectrum has been planned for high-frequency programs ranging from 8 to 12 GHz.

2.3. Design of metamaterial antenna on the body tissue in the hyperthermia application

At this step, three layer of body tissue including skin, fat and muscle in both simple (no metamaterial) mode, and then with metamaterial cover, is posited on the microstrip patch antenna. Figure 5 shows a metamaterial antenna with tissue.

In Figure 5, a skin layer of 1 mm thickness, a fat layer of 10 mm thickness, and a muscle layer of 40 mm thickness at a distance of Ag=25 mm from the microstrip patch antenna - which is actually an air gap - have been placed. Our aim is to investigate the effect of this air distance on the amount of output SAR. For this purpose, we have placed metamaterial coating at different distances from the tissue (Ag-hm) and then measured the amount of output SAR. As it is shown in Figure 5, by changing the Ag-hm distance, the gap between the metamaterial cover and the patch antenna (hm) will also change. Because, as mentioned, the overall air gap between the antenna and the tissue must remain constant (25mm).

3. Results and discussion

Figure 6 displays the s-parameter features of the metamaterial-free antenna.

It can be understood that the resonance frequency exists at about 9.63 GHz and that the S_{11} at this frequency is roughly -41 dB which has a reasonable value.

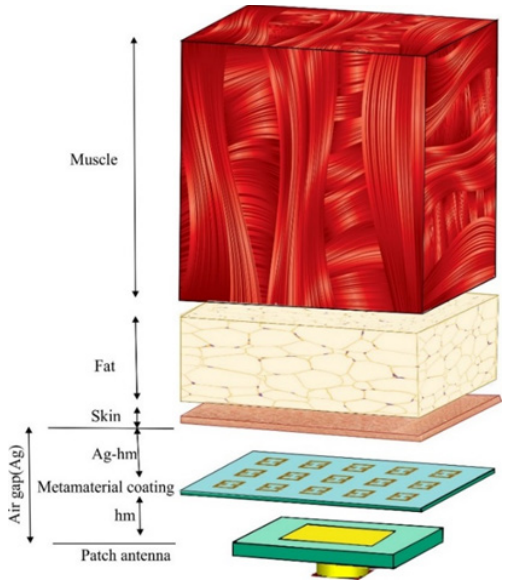


Figure 5: Geometry of the presented metamaterial antenna with body tissue.

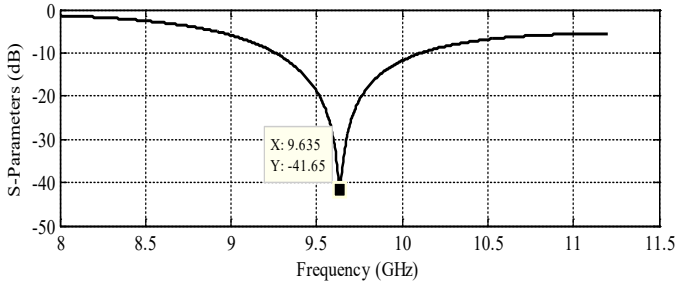


Figure 6: The antenna return losses without the use of metamaterial cover

The value of S_{11} in Figure 6, refers to the most power being transmitted to the antenna, and that is one of the fundamental definitions in the subjects of antennas. Figure 7 displays the pattern of radiation and the far field in metamaterial-free antenna. As shown in Figure 7, the gain is 7.66 dB at a frequency of 9.6 GHz. Half power beamwidth (HPBW) is the angle between two main beam points, 3-dB below the maximum power. That is 78.9 degrees here, and the main lobe direction is zero. Reducing the HPBW value is our aim for high-frequency applications. Because it indicates a rational directivity for our antenna [1].

3.1. Effect of airspace between antenna and metamaterial

In this section, there is not any gap between the metamaterial coverage and the patch of antenna, and the metamaterial coating including 15 SRR cells printed on a Teflon sheet has been directly mounted on it. In this case, Figure 8 displays the return losses of the antenna, and Figure 9 illustrates the return losses when 10.53 mm of air distance has been applied to the system.

In this situation, frequency has shifted roughly 0.7 GHz than the antenna without metamaterial. The metamaterial coverage is thus positioned at an optimal air gap of 10.53 mm as better results have been derived from the presence of airspace. According to Figure 9, the return loss of the patch antenna is about -50 dB due to the metamaterial coating structure which grants a remarkable performance. Also, the resonance frequency did not change substantially than the base antenna.

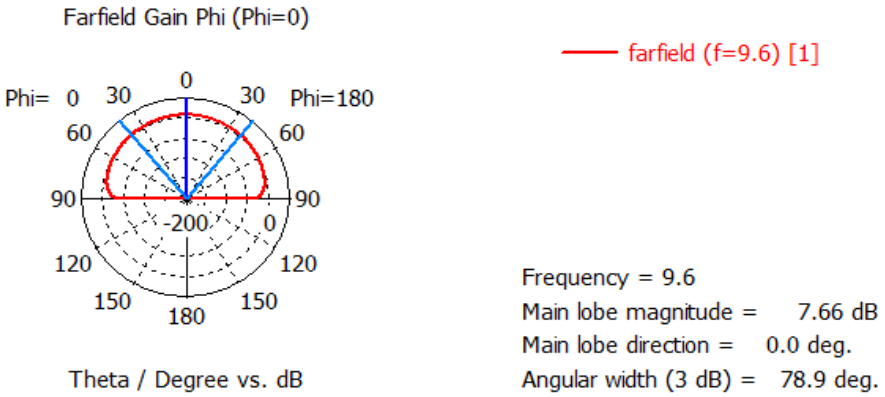


Figure 7: The radiation pattern of the metamaterial-free antenna

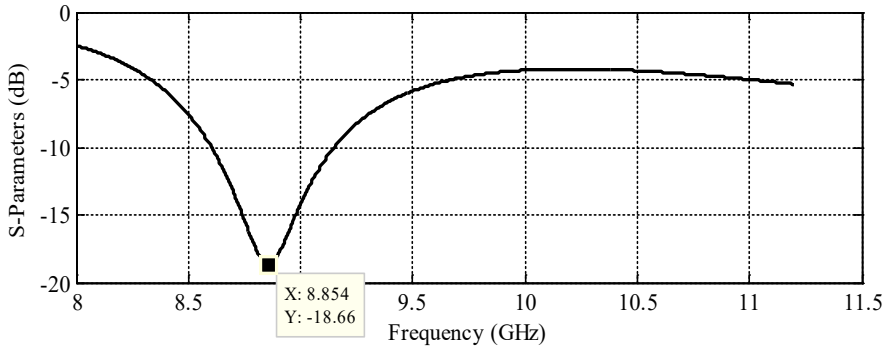


Figure 8: The return losses of antenna without airgap

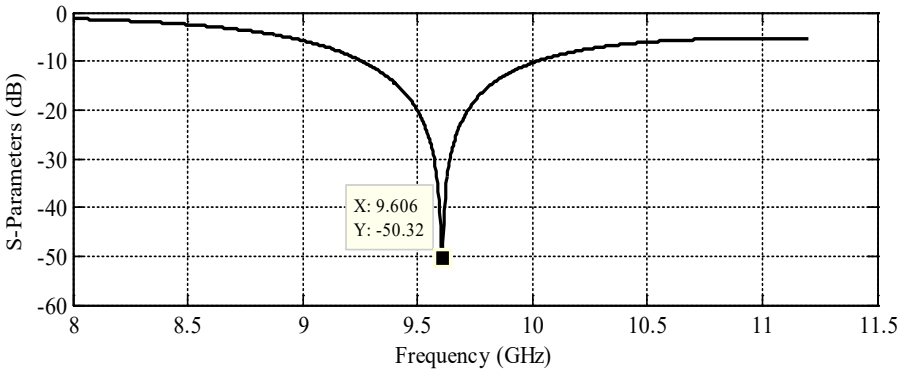


Figure 9: The return losses with airspace of 10.53 mm between the metamaterial and the antenna

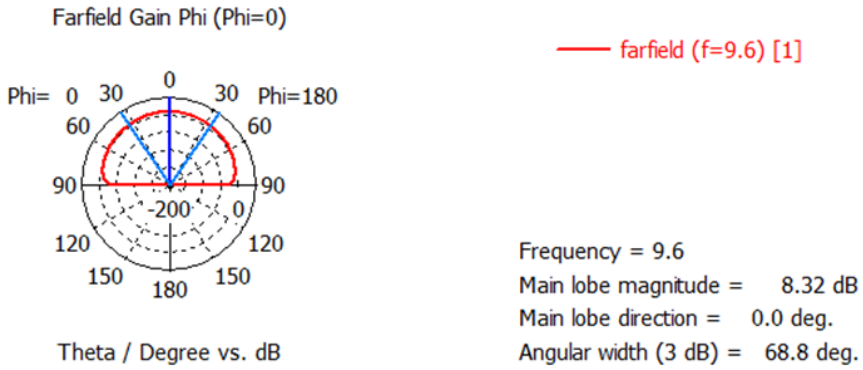


Figure 10: The radiation pattern of the metamaterial-coated antenna

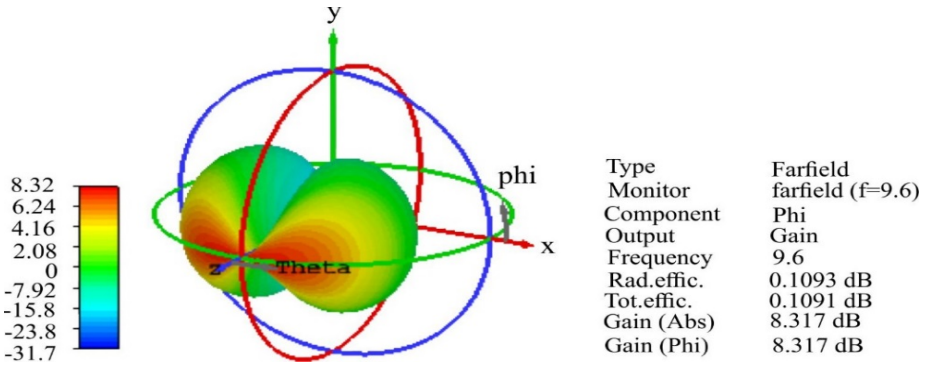


Figure 11: Three-dimensional view of the radiation pattern

Table 2: Comparison of outcomes between antenna without and with metamaterial (MM)

Parameters	MM-free Antenna	MM-coated Antenna
Return losses (dB)	−41.65	−50.31
Gain (dB)	7.66	8.32
HPBW (3dB)	78.9°	68.8°

Figure 10 illustrates the antenna radiation pattern with the metamaterial cover. As it can be observed in the Figure 10, the gain obtained is 8.32 dB at the frequency of resonance, the HPBW is 68.8 degrees and the main lobe orientation is zero.

Figure 11 indicates the radiation profile of the metamaterial-coated antenna in view of three dimensions. For comparing and better understanding, antenna outcomes in both types, with and without metamaterial, have been listed in Table 2. The beneficial impact of the usage of metamaterial cover in the function of the antenna is evident, according to Table 2.

As can be seen, gain has improved around 8.6 percent on the microscope patch antenna by supporting this designed structure. HPBW has also significantly decreased by about 12.8 percent. Such angle change has led to a significant improvement in the direction and a better centralization of the main beam of radiation pattern. In most of the high frequency applications, the more the antenna radiation can be transferred in the direction defined and intended or be oriented, the more the antenna would be preferable. Furthermore, in the frequency of resonance the S_{11} value also has a desirable change in this structure due to the metamaterial antenna configuration and indicates a suitable impedance matching [1].

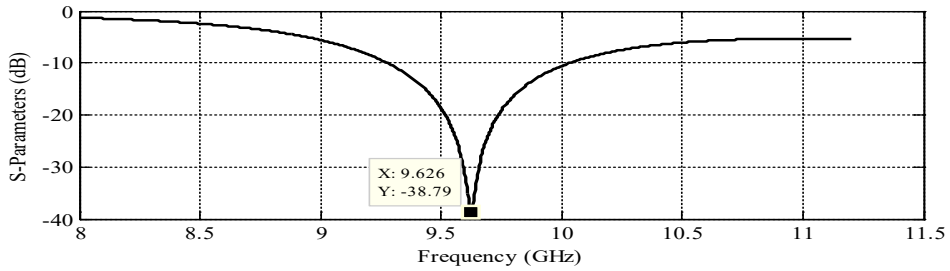


Figure 12: The return losses in the two-layer SRR mode

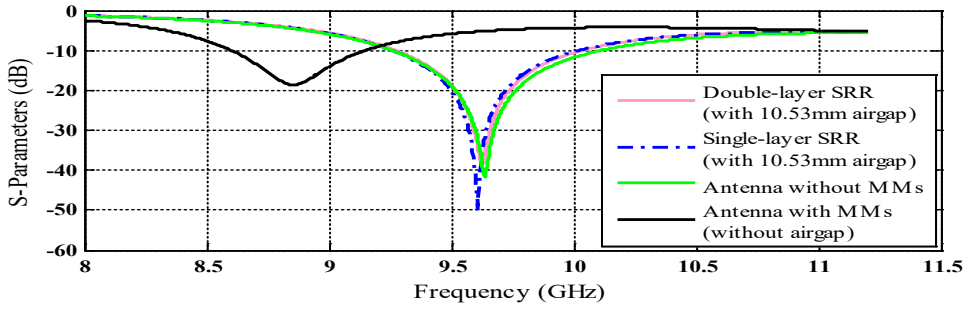


Figure 13: A comparison of the return losses and resonance frequencies of four mentioned antennas

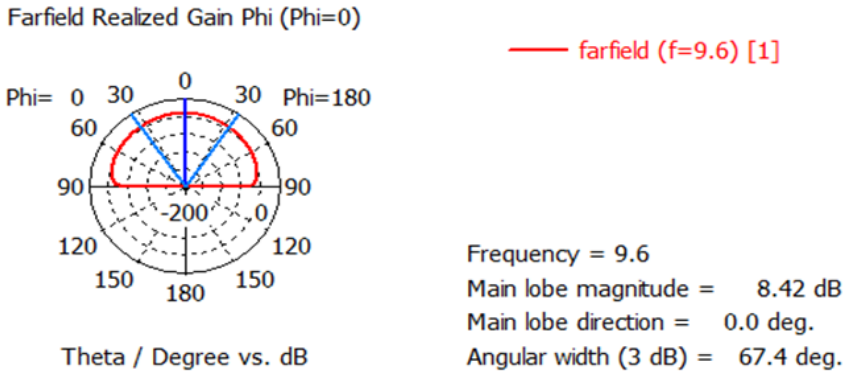


Figure 14: The pattern of radiation in double-layer SRR

3.2. Double-layer SRR effect

The metamaterial cover in this section comprises of two coats of SRR on the lower and upper plane of Teflon.

Because there are 15 SRR unit cells in each Teflon surface, the total number will be 30 unit cells. It has been positioned in the presumed airspace afterwards and outcomes have been compared to the monolayer SRR.

Figures 12 and 14 display the results of return losses and radiation pattern respectively. Also in order to comparing and gaining a better perception of the function of four mentioned types of antennas, results of the return losses have been represented in Figure 13.

The analogy of results between the single-layer SRR and the double-layer SRR can be seen in Table 3.

Table 3: Comparison of modeled SRR values in single-layer mode with double-layer mode

Parameters	One layer of SRR	Two layers of SRR
Return loss (dB)	−50.31	−38.78
Gain (dB)	8.32	8.42
HPBW (3dB)	68.8°	67.4°

As shown in Table 3, when the Teflon slab contains two upper and lower SRR layers, although the return loss has reached to -38, it would be in a highly admissible condition nevertheless, and the optimal output of the antenna is not harmed. The gain of antenna has had increasing pattern, plus HPBW has fallen by 1.4 degrees as well. Accordingly, in enhancing the radiation pattern, SRR in double-layer mode could become more efficient style [1]. In the current work, taking into account changes in antenna size and dielectric constant, the metamaterial-free antenna not only had stronger radiation than the compared article patch antenna in [19], but also some modifications in the form and quantity of the SRRs in the metamaterial antenna, as well as the amount of the air gap between the antenna and the cover, allowed noticeable improvements to be made. When we juxtapose the findings of metamaterial antennas in both Tables 1 and 3 (one-layer mode of SRR), we will find that the gain from present work has improved in comparison with the paper described. We decreased the HPBW by around 9.47 percent than the aforementioned article as well. Reducing HPBW implies that the centralization and power of the beam has improved [1]. For most of high-frequency antennas, that would be our key objective.

3.3. Effect of increasing the size and number of SRRs

At this stage, in order to modifying the previous operations, some changes have been applied to the shape, size, and number of unit cells. For this purpose, a piece of copper wire has been added between two squares. Also the number of unit cells has been increased from 5 * 3 unit to 5 * 4, and then to 5 * 5 unit, afterwards results have been investigated. New unit cell is shown in Figure 15. The results for the twenty and twenty-five resonators have been presented in the Table 4.

According to the Table 4, the radiation pattern has improved with the increase of size and number of SRR. But considering the value of S_{11} and comparing them in both twenty and 25 types, it is observed that as the number of resonators increases, the numerical value of S_{11} has changed. This means that return losses of antenna in metamaterial coating with 25 resonators have increased. Thus, it can be concluded that the increase in the number of resonators is faced with limitation, and if it exceeds a certain number, the antenna losses will increase.

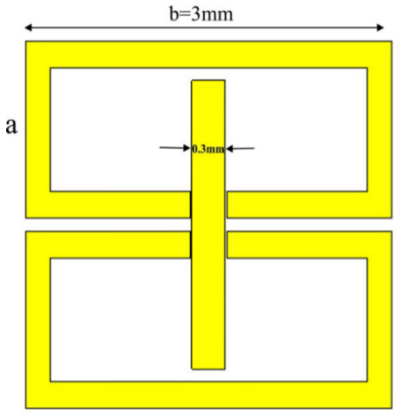


Figure 15: The modified unit cell by adding a piece of wire

Table 4: Comparison of the simulated results between metamaterial with 20 and 25 unit cells

Number of SRR	4*5 units	5*5 units
Frequency (GHz)	9.6448	9.6576
S_{11} (dB)	-23.66	-15.88
Gain (dB)	8.91	9.36
HPBW (deg)	60.2	55.9

But since in many space applications, antennas with sharp beamwidth are required for detecting and tracing purposes, it seems tempting to use this type of antennas as a lens. For this reason, it is not easy to ignore the sharpened beamwidth and high directivity, that is result of increasing the number of resonators. Therefore, to overcome this limitation, we have benefited a number of circular holes in the ground plane. The procedure is described in section 3.3.1.

3.3.1. Effect of inserting circular holes in the ground plane

The idea has been derived from the article [22]. The procedure of mentioned paper, has been to improve the antenna radiation pattern by using resonators along with creating nine holes in the ground plane. The use of circular holes, is an old method of limiting surface currents, which reduces the amount of return losses [22]. For this purpose, along with the use of metamaterial coating with 25 resonators, we have created nine circular holes with the radius of 1.5 mm on the ground plane. Figure 16 shows the metamaterial cover, and the holes created on the ground plane.

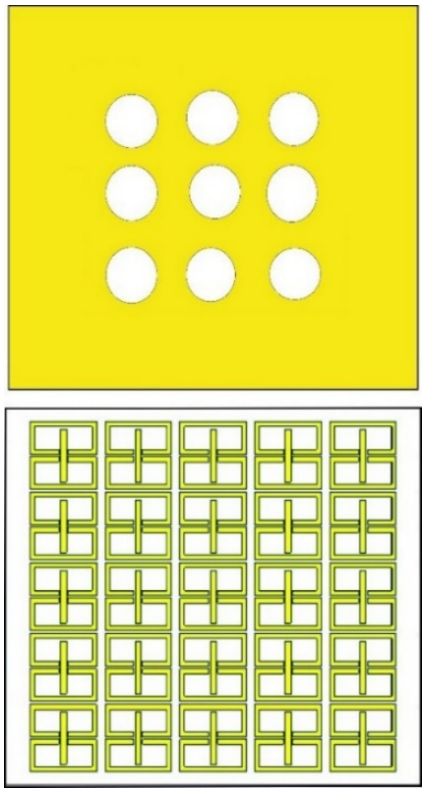


Figure 16: Circular holes created on the ground, and metamaterial coating with 25 unit cells

Table 5: Effect of different radii of holes on the return losses and resonance frequency

Radius (mm)	Frequency (GHz)	S_{11} (dB)
0.7	9.597	-15.74
0.8	9.635	-18.260
0.9	9.638	-19.94
1	9.635	-21.71
1.1	9.622	-24.29
1.2	9.603	-27.80
1.3	9.568	-29.24
1.4	9.539	-42.47
1.5	9.510	-38.46

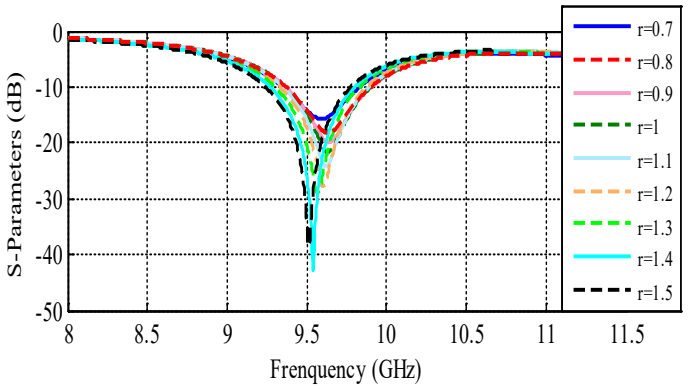


Figure 17: The return losses for different radii of circular holes

It should be noted that the radius of the holes can vary. The results of the effect of several different radii on return losses are presented in Table 5. Figure 17 also shows the return losses for the tabulated amounts derived from CST software.

As it is clear, the return losses have decreased dramatically as the holes' radii have increased. Frequency fluctuations are slightly visible as well. According to the Table 5, it can be understood that the hole creation significantly reduces the return losses and thus eliminates the antenna constraint caused by the increase of resonators.

It is worth noting that just because the frequency has also shifted slightly with the radius of the hole, in order to increase the accuracy of the radiation pattern results, we have extended the resonance frequency from 9.5 (for hole with radius of 1.5mm) to 9.6, which is the scope of this paper. This object can be easily achieved by altering the substrate dielectric constant from 2.2 to 2.1. For this purpose, we have replaced the Teflon of substrate with dielectric constant of 2.2 to PTFE existed in the software with dielectric constant of 2.1, in which case the frequency has reached 9.68. Also the amount of is about -32 which is quite acceptable.

Figure 18 demonstrates the radiation pattern of the antenna with the final improvement. Table 6 compares the results of this metamaterial antenna with those results in the [22].

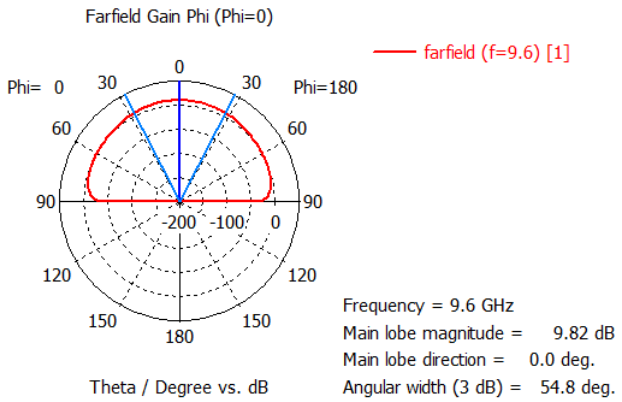


Figure 18: The radiation pattern of antenna including both 25 SRRs and 9 circular holes

Table 6: Comparison of simulated results between final improvement of this paper and paper [22]

antenna	Reference paper	Our work	Improvement (measured)
Return loss(dB)	-28	-32.78	-4.78 dB
Gain(dB)	9.6	9.82	0.22 dB
HPBW(deg)	68	54.8	19.5%

The creation of circular holes in the ground plane not only have significantly reduced the return losses, but in comparison with the Table 4 -when there is no hole- they have also increased the amount of antenna gain at a rate of 0.46 dB.

In general terms, the creation of the circular holes along with metamaterials consisting of large-sized and more-numbered resonators have impressively increased the efficiency of the antenna.

This method can therefore be regarded as a solution for reducing antenna losses and overcoming the antenna restriction due to the increase in resonator numbers.

3.4. Effect of Metamaterial layer on the patch antenna in hyperthermia application

Table 7 shows amounts of SAR obtained for different distances of Ag-hm. To better realization of the Table 7, results have been shown in Figure 19.

As it can be seen in the diagram, the lowest amount of output SAR is pertinent to 24 mm, namely maximum Ag-hm distance.

By reducing the distance between the metamaterial coating and the body tissue, the numerical value of SAR has increased relatively, so that at 6.25 mm, the amount of SAR has reached its maximum value of 15 w/kg.

It can be concluded from Table 7 that the distance between the tissue and the metamaterial antenna plays a significant role in the output SAR value. In other words, the SAR values required for different usages can be obtained by varying the distance.

In the next phase of the experiment, the hotspot of the lowest SAR amount obtained (5.2 w/kg), at the maximum distance between tissue and metamaterial, was estimated at approximately 4.8 w/kg. In the following, the rest of the SAR simulations values were normalized accordingly.

A few examples of the results in the normalized SAR value are given in Figure 20.

Table 7: SAR amounts obtained for different distances between metamaterial coating and body tissue

Air Distance [Ag-hm] (mm)	SAR Amount (w/kg)
24	5.2
20	9.25
18.75	8.99
16.66	10.1
15	10.7
12.5	11.25
10	12.6
8.34	13.9
6.25	15
5	14.5
1	10.1

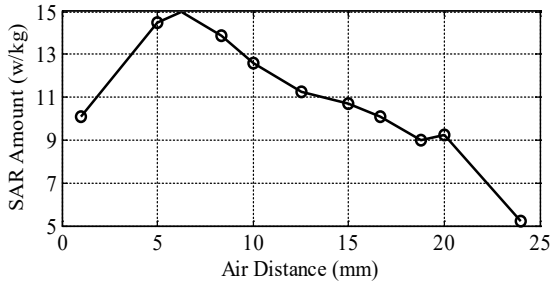


Figure 19: Schematic view of SAR amounts based on different gaps delineated with Table 7 data.

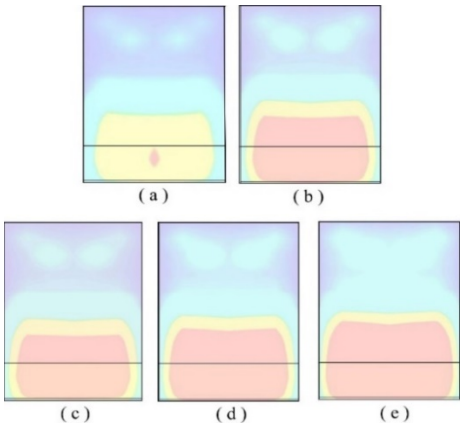


Figure 20: Normalized SAR based on Minimum SAR (in 4.8 w/kg). (a) Minimum SAR with value of 5.2 w/kg in distance of 24mm. (b) SAR amount: 9.25 w/kg in air gap of 20mm. (c) SAR amount: 8.99 w/kg in air gap of 18.75mm. (d) SAR amount: 10.7 w/kg in air gap of 15mm. (e) SAR amount: 13.9 w/kg in air gap of 8.34mm

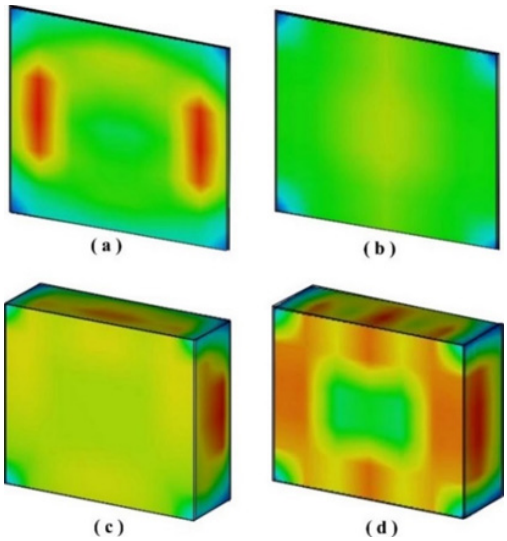


Figure 21: Comparison of heat permeation between usual and MMs antennas in two tissue layers of skin and fat. (a) Skin layer in usual simple antenna. (b) Skin layer in MMs antenna. (c) Fat layer in usual antenna (d) Fat layer in MMs antenna.

As it can be observed in Figure 20, with increasing the amount of output SAR, a larger area of tissue is also involved in heat. Therefore, with altering in this distance we will be able to increase or decrease the amount of the output SAR. To describe precisely, it will be feasible that the amount of heat required come under controlled to some extent.

The use of metamaterial coating on the antenna has another advantage, which is remarkable in its form, and that is the conduction of electromagnetic energy and the heat to the higher depths of the body tissue. This subject causes to reduce damages due to heat in long time to the surface tissues.
Figure 21 shows the two superficial layers of tissue, skin and fat, that compare the amount of transition heat between both types of simple microstrip antennas and metamaterial antennas.

As it is noticeable in Figure 21, in the usual antenna, the first surface layer of tissue namely skin has been burnt and the rest of heat has penetrated to the initial layers of fat, whereas in the metamaterial antenna, heat has penetrated into the deeper layers of fat and there is no burn on the skin tissue.

4. Conclusion

In this study, we took advantage of a type of metamaterial formation as a coating layer upon a microstrip patch antenna with coaxial feeding to upgrade the efficiency of antenna. The proffered small-size antenna was modeled in a resonance frequency of 9.6 GHz, whereby the metamaterial structure layer was used to achieve a highly desirable orientation. The gain of the antenna increased significantly as well, and the half-power beamwidth was decreased by more than 20 degrees than the normal patch antennas, which tend to be striking in spatial tracing applications considering the ordinary quantities of beamwidth in the usual antennas.
Using of metamaterials in antennas not only causes to reduce losses amount, increase the gain and sharpen the radiation as a lens in satellite usages, but also can alter the depth of penetration in medical applications, as well as acquire various SAR amounts, by adjusting the proper distance. In fact, more degrees of freedom will emerge for different applications. Specifically, by using different shapes of unit cells, other forms of these designs will present different outcomes.

References

[1] Jahangiri, Mahsa, and Saman Rajebi. "Effect of square srr structure on the radiation pattern of the patch antenna in the x-band." *2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI)*. IEEE.
[2] VESELAGO, Viktor G. The electrodynamics of substances with simultaneously negative values of and μ. *Soviet physics uspekhi*, 1968, 10.4: 509.
[3] PENDRY, John B., et al. Magnetism from conductors and enhanced nonlinear phenomena. *IEEE transactions on microwave theory and techniques*, 1999, 47.11: 2075-2084.
[4] SMITH, David R., et al. Composite medium with simultaneously negative permeability and permittivity. *Physical review letters*, 2000, 84.18: 4184.
[5] ISLAM, Mohammad Tariqul, et al. A negative index metamaterial-inspired UWB antenna with an integration of complementary SRR and CLS unit cells for microwave imaging sensor applications. *Sensors*, 2015, 15.5: 11601-11627.
[6] ENGHETA, Nader. Metamaterials with negative permittivity and permeability: background, salient features, and new trends. In: *Microwave Symposium Digest, 2003 IEEE MTT-S International*. IEEE, 2003. p. 187-190.
[7] WITHAYACHUMNANKUL, Withawat; ABBOTT, Derek. Metamaterials in the terahertz regime. *IEEE Photonics Journal*, 2009, 1.2: 99-118.
[8] PATEL, B. D.; NARANG, Tanisha; JAIN, Shubhangi. Microstrip Patch Antenna-A Historical Perspective of the Development. In: *Conference on Advances in Communication and Control Systems*. 2013.
[9] CONSTANTINE, A. Balanis, et al. Antenna theory: analysis and design. *MICROSTRIP ANTENNAS, third edition, John wiley & sons*, 2005.
[10] RAHMAN, Md Mahfuzur. *Small size coupling feed and inductive shorting antenna for wide bandwidth, increased gain and efficiency with low specific absorption rate (SAR) operation*. 2016. PhD Thesis. Michigan Technological University.
[11] KARADE, Abha R.; ZADE, P. L. A miniaturized rectangular microstrip patch antenna using SSRR for WLAN applications. In: *Communications and Signal Processing (ICCSP), 2015 International Conference on*. IEEE, 2015. p. 1002-1004
[12] SMITH, D. R., et al. Electromagnetic parameter retrieval from inhomogeneous metamaterials. *Physical review E*, 2005, 71.3: 036617.
[13] HOSSAIN, M. I., et al. Low-SAR metamaterial-inspired printed monopole antenna. *Applied Physics A,* 2017, 123.1: 8.
[14] ALI, Tanweer; MOHAMMAD, Saadh AW; BIRADAR, Rajashekhar C. A novel metamaterial rectangular CSRR with pass band characteristics at 2.95 and 5.23 GHz. In: *Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2017 2nd IEEE International Conference on*. IEEE, 2017. p. 256-260.
[15] Nijhawan, Geeta, Siddharth Sagar Nijhawan, and Minu Sethi. "Hyperthermia Treatments." *Noble Metal-Metal Oxide Hybrid Nanoparticles*. Woodhead Publishing, 2019. 241-263
[16] Mendez, Hector Fabian Guarnizo, Mauricio Andrés Polochè Arango, and John Jairo Pantoja Acosta. "Hyperthermia Study in Breast Cancer Treatment." *Workshop on Engineering Applications*. Springer, Cham, 2018.
[17] Jaffar, Nabilah Abdul, et al. "An overview of metamaterials used in applicators in hyperthermia cancer treatment procedure." *2017 International Conference on Electrical, Electronics and System Engineering (ICEESE)*. IEEE, 2017.
[18] Ebrahimi-Ganjeh, Mohammad Ali, and Amir Reza Attari. "Study of water bolus effect on SAR penetration depth and effective field size for local hyperthermia." *progress in Electromagnetics Research* 4 (2008): 273-283.
[19] Paswan, Pradeep, et al. "Performance enhancement of coaxial feed microstrip patch antenna using left-handed metamaterial cover." *2014 IEEE Students' Conference on Electrical, Electronics and Computer Science*. IEEE, 2014.
[20] SEO, Je Hyung, et al. Distribution of THz electric field in the split-ring resonator metamaterials based on the thin film geometry. *Current Applied Physics*, 2016, 16.3: 329-334.
[21] GARG, Ramesh. *Microstrip antenna design handbook*. Artech house, 2001.
[22] Fang, Chao, and Feng Xu. "Performance enhancement of coaxial feed microstrip patch antenna using left-handed material and PBG circular holes." *2014 IEEE International Conference on Communiction Problem-solving*. IEEE, 2014.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 4, 308-313 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Enhancing Bodily Movements of the Visually Impaired Children by Airflow

Fang-Lin Chao[1,*], Hung-Chi Chu[2], Liza Lee[3]

[1]*Department of Industrial Design, Chaoyang University of Technology, Chaoyang University of Technology, 436, Taiwan R.O.C.*

[2]*Department of Information and Communication Engineering, Chaoyang University of Technology, 436, Taiwan R.O.C.*

[3]*Department of Early Childhood Development & Education, Chaoyang University of Technology, 436, Taiwan R.O.C.*

ARTICLE INFO

Article history:
Received: 22 June, 2019
Accepted: 02 August, 2019
Online: 07 August, 2019

Keywords:
Airflow
Interaction
Bodily movement
Visually impaired

ABSTRACT

This study involved three departments of teachers and students. The concept of the flow-motivated exercise implemented with a set of flow nozzles. A computer-controlled airflow with specific music introduced, and the motion responses from the participants were analyzed during the initial test and field study. During the initial examination, we adjusted the magnitude and distance setting of the flow and allowed the children to grab the air without music. Structural design, magnetic valve, and interface were implemented based on the insights from the initial test. The redesign performed in a sitting position to reduce the need for students to deviate from the location and fit for physical disabilities during exercise. Four times of visually impaired students were arranged to observe the changes in the movements of students after being guided by the airflow during the teaching process. The performance of the fourth week is better than the previous three weeks. Almost all blind children can make movements in the right position. The cross-domain cooperation experience was beneficial to the facility design.

1. Introduction

Most of the visually impaired education is established using state funds, and many subjects are integrated with the existing school to incorporate them into society. There are not many schools for the visually impaired students in Taiwan, and some children attend the school for the mainstream setting. Under the guidance of special education teachers, the content of the textbooks for visually impaired students is designed the same as the content for mainstream students.

The core curriculum includes compensatory or functional academic skills, orientation and mobility, social interaction, independent living skills, recreation and leisure skills, technology, and visual efficiency skills [1]. Some activities provided independent in dealing with his environment at a residential school. The students with visual impairment cannot comprehend the entire picture by extracting a lot of information. Visual impairment has complex experience with the spatial world. Study [2] involving fieldwork in everyday life situations to gain insight of their experienced. The allocation of limited instructional hours and the shortage of qualified specialists are significant factors that influence effective instruction.

*Corresponding Author: Fang-Lin Chao, flin@cyut.edu.tw

Experienced teachers are scarce; thus, teaching aid facilities have to be developed on an immediate basis. Teaching aids reduce the problem of repeated guidance and allow students to obtain knowledge in a safe environment. Specialized schools in Taiwan for the visually impaired accommodate students with multiple barriers, such as slow limb movements and cerebral palsy. We observed the activities of these students after class. The students do not have much language communication with each other. When they are bored, they lie on the table or unconsciously tamper with the objects around them. Therefore, the involvement of the management is required. Once the rhythm teaching aids are placed on the table, they are engrossed in playing with the device and have a will to explore (Figure 1). The major mode of teaching is oral instructions. Some students cannot keep up with the progress when they are not attentive. They can explore and feel based on the teacher's instructions by using teaching aids. Rhythm and music-related teaching aids help students, especially in terms of instrument appearance- and rhythm-related learning. Learning through teaching aids, which cannot be easily expressed in words, presented its unique benefits.

1.1. Physical education

The physical education classes of most students include competitive events, such as running, long jump, or ball-related

games. However, these types of games do not help in enhancing the skills required by the visually impaired students for living. Physical movement and rhythm are necessary skills for the visually impaired students to keep them healthy and to help them obtain higher self-confidence.



Figure 1: Behavior observation of visually impaired children in an ordinary kindergarten classroom with their teacher [3].

The performance of the visually impaired individuals in physical education is directly associated with the severity of visual impairment, activities performed, the ability to ambulate, and family attitudes [4]. Postmenopausal changes significantly decrease the body's mechanisms for maintaining postural stability. Another study [5] analyzes the effect of sensorimotor exercise on changes in postural stability of visually impaired women. "Late blind persons had a slower walking speed, shorter stride length, and the longer time of gait" [5]. Their gait represents an attempt to adapt to maintain a stable posture during walking.

We must know the students' characteristics (limitations, abilities, and needs) in physical education programs. A school should also be aware of the teaching burden. Visually impaired children face a barrier in being integrated into the general physical education classes [6]. These children display lower physical activity levels because of the lack of instruction and practice.

Teachers sometimes lack sufficient knowledge pertaining to the method of including students. Inadequate preparation, lack of appropriate equipment and curricula, and insufficient teaching hours caused dominant barriers. Consequently, children demonstrated delays in reaching developmental milestones. A sample of 125 visually impaired children was observed [7]. The self-personality and prosocial behavior of the children were explored with respect to their mobility training. The personality well-being and social ability are correlated with the ability to be mobile.

1.2. Actuator and sensor

Technology-assisted learning has been a trend in the recent years. People use many apps to explore data, and visually impaired students can expand the scope of activities and overcome physiological limitations through the assistive technology. A study [8] developed smart connected devices for playfulness by using littleBits and Loaded Dice. Authors conducted experiments using visually impaired students and engaged in playful exploration and investigated the extent of providing imaginative design concepts. Another study [9] provided a visually impaired user interface by using embossed two-dimensional patterns that can be read using haptic sensing. In [10], a researcher introduced a haptic device that can notice the presence of obstacles inside an environment. A prototype that was implemented using a short cane and an embedded smart sensing strategy was presented.

The engineers used senses other than the visual sense to transmit a message to the visually impaired. For example, tactile pressure, temperature field, airflow field, and sound can be reconsidered for such applications that are focused on the visually impaired people. When these applications are aimed at the visually impaired population, we must reduce the cost and complexity of the technology because a high price renders the technology unaffordable and a complicated technology cannot be used reliably.

Kraus [11] designed teaching tools that are appropriate for exploring the solar system and different astronomical subjects. The magnetic field, temperature drop, and airflow connector presented physical space for the visually impaired people. The availability of the tactile constellation presents apparent magnitudes and displays the star distribution in the Cassiopeia constellation.

1.3. Flow-Motivated Interactions

This study focuses on three considerations: (1) providing haptic attention of airflow, (2) encouraging natural response from children, and (3) engaging to relate the body motion based on the excitation. Although the target users are visually impaired children, ordinary people or the elderly can follow the rhythm and find the pattern to play. The flow-motivated game uses a set of flow nozzles that can generate segmented air flow through a compressor. An air compressor is readily available in a shop. The compressor increases the pressure in the air-intake system [12] and generates a mass flow. The intensity of the mass flow can be sensed by the skin.

The airflow should be sensed by the skin; however, a very strong airflow causes discomfort. A small pipe sends the jet stream, and the high-speed air stream simultaneously causes a concomitant acoustic field with a sharp sound. Moreover, the time required for a child to respond to the airflow by performing actual body action should be considered. The response to the airflow varies from person to person. Some people move in the direction of the airflow, and some people grab the source of the airflow. This design encourages a spontaneous response and does not force the user to respond in specific manner. When a user is not pressurized to act in a specific manner, he or she can easily respond spontaneously.

The third design consideration was engagement to relate the body motion based on the excitation. The stimulation of the airflow cannot be too frequent and unpredictable and within a reasonable reaction time. The rhythmic stimuli are matched with the rhythm of the tones, thus promptly providing airflow guidance while providing the advice using the sounds. We provided four–six airflow spouts at appropriate body parts. The upper body and the lower body have three spouts each. The airflow field of each spout could be felt and was used based on the position of the person standing.

2. Materials and methods

At initial test, we used four impingement spouts without music. Teacher controls the gas flow and let the children try to grab the air. The four heads are connected to the air compressor respectively, and we control the opening and closing of the airflow by hand. We need to understand the impact of the height, position, and distance of the nozzle on the child before design. Teacher fixed the spout with a height-adjustable bracket, which was initially located at the waist of the children.

During test the position was adjusted to avoid being touched. Students listen to the teacher's introduction before they start. We

found that at the beginning, they were a bit nervous and the action was not very casual. The mentality may be a bit like facing the exam, not in a state of play.

2.1. Initial test

The experimental design used four of airflow without music. Most subjects stood still in the center without body movement and facial expression at the beginning (Figure 2).

When the airflow came out from the four ventilators at the same time, subject A,C,E,F,G,K,L,M,N,O didn't move at all.

They stood still until the air stopped. (O-A,C,E,F,G K,L,M,N,O -0325-1,2,3,4) (O-A, C, E, F, G, K,L, M, N, O-0325-R)

When subject K, L, M, N, O felt the airflow, they didn't know how to move because they are too young.(I-0325-T)

A few subjects didn't do any movement but did smile. Subject B and D didn't do any bodily movement but a few seconds later, they started smiling. (O-B, D -0325-1, 2, 3, 4) (O-B, D -0325-R)

A few subjects didn't show any expression, but moved their hands. Subject B and D didn't do any bodily movement but a few seconds later, they started smiling. (O-H, I, J -0325-1, 2, 3, 4) (O-H, I, J -0325-R)



Figure 2: Initial test setup for evaluating how visually impaired children respond to gas spreading

2.2. Discussion of initial test

This test revealed some problems.

(1) Some children did not know how to respond and became frustrated. When the students asked "Why am I coming?", they intended to tell the teacher "what should I do in this situation?". Formal learning habits constrain the mood of participation.

(2) Although some children were attracted to the flow in the first few rounds, they became bored because the same motions were repeated. Simple physical movements allowed them to respond only mechanically. The lack of self-inherent motivation reduces their interest.

(3) Some children paid so much attention to others' opinions that they became distracted and began playing. The students should be integrated in the event with a relaxed mood.

(4) The airflow produces a sharp sound. The air flow must be moderate. The motor generated considerable noise when the air pressure of the air compressor reduced. Children can get scared because of the sudden changes in the outside world.

(5) We determined whether a garment blocks airflow and found that the students did not rely solely on the airflow sprayed on the skin. When the airflow ejected, the turbulence air generated a sharp sound. Visually impaired children are sensitive to hearing, and they can determine the source of the airflow based on the noise.

Design of Flow Control

For realizing a simplistic design and the lowest price for the visually impaired, we avoided the complicated design to reduce possible failures. A central computer with touch panel controls the progress of the song and direction of the gas. Figure 3 displays the structure of the system. In the concept proposal, we suggested that the ground has a pressure sensor to determine the user's position and that infrared sensors can be installed on the bracket to know how the user moves.

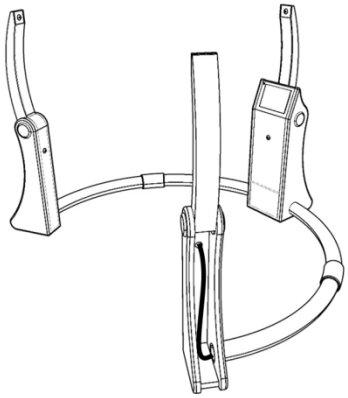


Figure 3: Diagram of the flow-motivated game with a magnetically controlled valve. Based on the problems found from the test model, a setup model with three nozzle arms was selected.

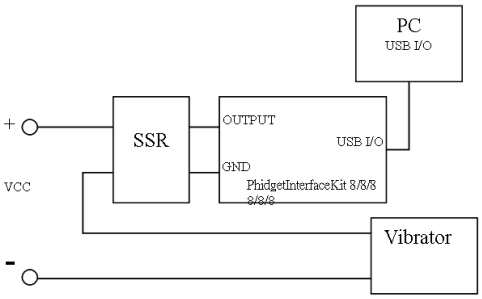
The initial test model was arranged with an air compressor and four nozzles. When the air compressor was switched on, the air reservoir of the air compressor was gradually pressurized. At this time, the control valve was opened to eject air. We used a magnetic valve to control the air outlet switch, which can be controlled remotely using an electrical signal.

The height of the three brackets was set based on the height of the children, and part of the gas was transmitted from above. Thus, the height of the bracket was set at 160 cm (Figure 3). There is an air outlet at the upper and middle sections of the bracket. The central part contains a rotating shaft, and the top arm can be folded into the lower feet. The circular frame in series was placed on ground to stabilize the three main arms. Control wires and airflow conductors were routed from the central computer through those tubes.

2.3. Circuit Design

Prototypes involving the control circuit and user interfacing software were developed. A USB interface control panel and solid-state relay (SSR) were placed on the circuit board (Figure 4). The Phidget Interface Kit [13] was utilized intangible user interfaces; the interface kit screen controller provides accurate results. We used the Phidget Interface 8/8/8 input–output board servers as a communication channel with the host computer.

The program controls the switching of the gas nozzles by using a preset sequence. The USB interface board cannot provide sufficient electrical current to switch the magnetically controlled valve. We constructed an external power board to ensure adequate current supply. An SSR has neither moving parts nor arcing contacts, which are often the primary cause of failure [14]. This provides a high degree of reliability and a long service life.



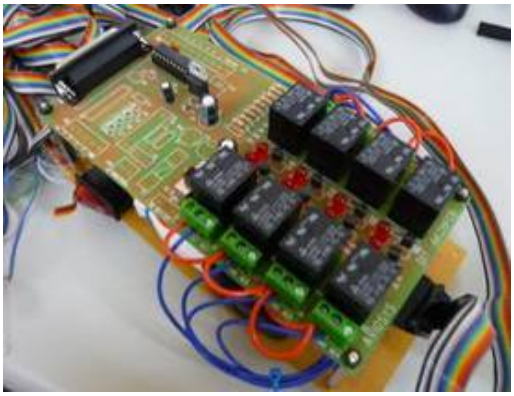


Figure 4: Circuit diagram and SSR circuit board with an external power supply [14].

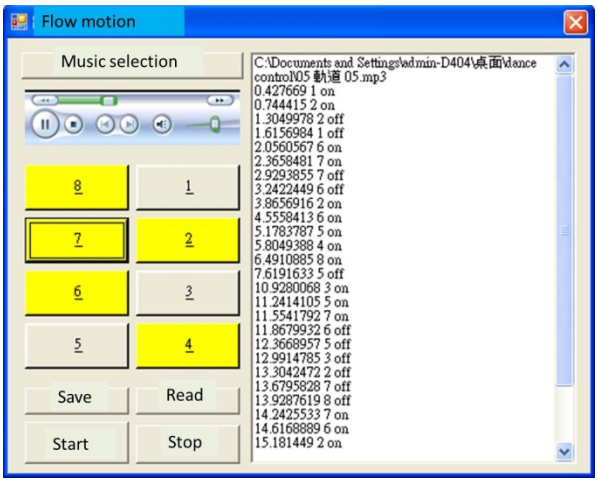


Figure 5: Software interface layout design with a sequential setting of the gas outlet and music control. The yellow blocks indicate active channels.

2.4. Dancing control editor and graphical user interface

As displayed in Figure 5, the interface control program has a graphical user interface that was developed on Microsoft Visual Studio 2005. The graphical user interface displays the current state of each switch. The button with the yellow background color indicates that it is currently activated. The text area on the right-hand side lists the sequence of the switching actions. A user can manually control the switching by directly clicking on the buttons at the specific time while the music is playing.

The status indicator, file read, and start–stop buttons are above the music editor. We can see the blocks 8, 7, 6, 2, and 4 in yellow, which implies that five nozzles are switched on. In the subsequent tests, we found that when most of the spouts simultaneously ejected, the students got confused while determining the source.

2.5. Multidisciplinary

This study comprises three departments of teachers and students. The Department of Industrial Design is responsible for proposing the structure of the concept. The students of the computer communication department are familiar with the circuits and programs and can construct the control circuit and the software to develop a working model. The students from the Department of Early Childhood Education invited children to try out prototypes in special education schools and found the areas of improvement in trial teaching.

In the past, industrial design paid more attention to the performance of modeling pursuit; computer students' interest in the pursuit of technology, they often prefer multifunctional betting into products. In the course of teaching, we understood that the function of a product is not the main part, and the user is the main part. A teacher has to adopt the story-telling method in the guiding process to allow the users to accept the taught content wholeheartedly. The experience of integrated research can adjust the design parameters and get an appropriate product design.

3. Field Study Results

The purpose of the field study was to examine a combination of music and airflow to enhance the bodily movement of visually impaired people. The methodology of this study was based on a qualitative study using interviews with students and the classroom teacher and observation reports from four observers and the researcher. Fifteen 4–19-year-old visually impaired students who enrolled in a school for the visually impaired in Central Taiwan were selected through purposive sampling for the study.

3.1. Materials and methods

The assessment instruments included structured observation forms to gather data on the reaction of the subjects from the researcher and four observers who were trained graduate students, interview reports from the teacher and the subjects at school. All interview sessions were recorded on video, and these were viewed and reported by four observers. The following is the coding symbol of the study:

Table 1: Coding symbol for music and flow activities

Coding	Meaning
O-A-0325-4	Observation of subject A on March 25, 2016 by observer 4
O-B-0325-R	Observation of subject B on March 25, 2016 by the researcher
I- 0325-T	Interview with the teacher on March 25, 2016
I-C-0325-R	Interview with subject C on March 25, 2016 by the researcher

The experimental design utilized four ventilators of airflow simultaneously with music. While revised experimental design from airflow only to play with music simultaneously, some of the subjects smiled but did not do too much movement, especially the five preschoolers. Most subjects' bodily movement was hand movement only.



Figure 6: Children interacting with the airflow

3.2. Results

In order to facilitate identification, we use the English symbol A B C D... to represent the participating students. While subject A, B, C, D, E, F, G, H, I and J listened to the music and felt the airflow, they didn't have any bodily movement but a few seconds later, they started smiling. (O- A, B, C, D, E, F, G, H, I, J-0415-1, 2, 3, 4) (O-A, B, C, D, E, F, G, H, I, J -0415-R)

The younger subject K said "It's music" while listening to the music. He also used his hands to feel the airflow. (O-K-0415-1, 2, 3, 4) (O-K -0415-R)

The subjects L, M, N and O all moved their hands to feel the airflow while the music played. (O-L, M, N,O-0415-1, 2, 3, 4) (O-L, M, N, O -0415-R)

When the researcher asked the subject K, L, M, N, O "Do you like music?" they all answered "Like." (O-K, L, M, N,O-0415-1, 2, 3, 4) (O- K, L, M, N, O -0415-R)

The subject K smiled. He said he liked the music and airflow. (O-K-0415-1, 2, 3, 4) (O-K -0415-R)

Due to the limited movement, the study changed the dynamics of airflow and music. Some subjects did not like the ventilators with strong airflow simultaneously with loud music. Some subjects said it was hard to feel anything from the four ventilators with weak airflow simultaneously with soft music.

When the four switches of airflow were opened, the researcher asked subject K "Would you like to move?" He said "Yes." But subject K needed to be held with the researcher. He wanted to find the airflow with the researcher. (O-K-0415-1, 2, 3, 4) (O-K -0415-R)

When the four switches of airflow were opened, subject L moved his body to face the air and moved his hands to touch the tube, (O-L-0415-1, 2, 3, 4) (O-L -0415-R)

When the four switches of airflow were opened, subject M stuck out his tongue and opened his mouth to feel the air. (O-M-0415-1, 2, 3, 4) (O-M -0415-R)

During four ventilators of weak airflow simultaneously with soft music, the subject O liked it and smiled. (O-O-0415-1, 2, 3, 4) (O-O -0415-R)

When the researcher asked the subject P "Do you like the air and music?" he said "like." (O-P-0415-1, 2, 3, 4) (O-P -0415-R)

When the ventilators were opened randomly, most older subjects could find the direction to adjust and move their hands and feet to feel it. (O- A, B, C, D, E, F, G, H, I, J -0415-1, 2, 3, 4) (O-A, B, C, D, E, F, G, H, I, J -0415-R)

3.3. Discussion of field test

The visually impaired children like music; thus, music it could be a good motivator for them to exercise. The study also found that children enjoy catching the flow. Sometimes their hand touched the air outlet, and this disturbed the original setup of the gas flow direction (Figure 6). This free exploration usually occurred in the first class and should be encouraged. With the teacher's guidance, the students gradually responded to such airflow with their body. Giving the students freedom to explore is also an essential factor. When students reflect based on specific rules, they lose their interest. Moreover, the interaction between peers affects an individual's attitudes. Peers' experience of trying can also stimulate the willingness of the students to try.

The results suggest avoiding frequent changes in the direction of the airflow. Students give up when they are unable to respond appropriately. The design was such that the students could be guided using a small amount of airflow from above or below depending on the characteristics of the music. Students become familiar with it by gradual learning.

The sound of the motor is loud when the air compressor is on, which scares the children. One of the solutions is to move the air compressor to another space or use a small air compressor. Then, the air volume ejected by the device can be reduced once the students are familiar with the device.

4. Redesign

The redesign adopted sitting position (Figure 7) to reduce the need for students to deviate from the location and fit for physical disabilities during exercise. The air outlet is located at the lower edge of the table to avoid being deflected by the student. Infrared sensors arranged on the side of the air outlet end, and a variable resistor provided for adjusting the distance of infrared emission and reception. Visually impaired students with physical disabilities sit in the middle and shift their bodies with the direction of the airflow; this prevents them from falling due to instability during the shift. Four times of visually impaired students were arranged to observe the changes in the movements of students after being guided by the airflow during the teaching process. There were six subjects under test, and Professor Li was composing music for visually impaired children. The TA will fill in the average performance (1-5 points), range of action, and emotional response of the children.

The first week found that due to the experience of no contact airflow in daily life, there was rejection at the beginning of the test. Subject A needs the teacher to encourage him to find the outlet point actively. Subject B can actively touch after the teacher's

guidance, but the body does not move when the air jets. After the guides explained, they were able to proactively find the airflow, but the movements did not change much (Figure 7a). The performance of the fourth week is better than the previous three weeks. Almost all blind children can make movements in the right position (Figure 7b) and can feel the way the airflow interacts. The test results show that the overall average performance is increasing from 1.6 to 4.3. The average number of follow-up airflows was 4.31. In the magnitude of the action (can swing the limb) up to 3.12. When the music and the rhythm of the airflow match, it is more inspiring to the children's movements.



(a) (b)

Figure 7: Response of the respondent (a) First week, (b) the Fourth week - Encourage students to respond with a silk scarf.

5. Conclusions

We presented a flow-motivated facility that adopts noncontact techniques for enhancing the exercising behavior. The airflow reinforces active participation of visually impaired children and increases positive interactions. The results of the study revealed that visually impaired children like music; thus, music could be a good motivator for them to exercise. They were beneficial to develop more holistic equipment for the visually impaired. Brief comparisons of the results against previous findings in [4] and [6], the decreasing of the body's mechanisms for maintaining postural stability can gradually improve after systematic learning and stimulation. Children enjoy catching the airflow; thus, sometimes their hand touched the air outlet and disturbed the original setup of gas flow direction. The study comprised three departments of teachers and students. This cross-domain cooperation experience was quite rare for us. Appropriate guidance during the teaching process is vital for the user to accept the product. The integrated research can adjust the design parameters and obtain a proper product design. The use of appropriate technology solves user problems and reduces the user's frustration.

References

[1] Hatlen, P., "The core curriculum for blind and visually impaired students, including those with additional disabilities" Review 28(1), 25-32, 1996.

[2] Cecilie H. Anvik, "Embodied spaces in the making: visually impaired people, bodies and surroundings" Scandinavian Journal of Disability Research, 11(2), 145-157, 2009. DOI: 10.1080/15017410902830710

[3] Chao, F.L., Lee, L., Chu, H.C. & Wu, W.T., "A Pilot Study on Applying Combination of Music and Airflow to Enhance Bodily Movement of Visually Impaired" in P. Resta (Ed.), Proceedings of Society for Information Technology & Teacher Education International Conference, 3564-3569, Austin, Texas, USA, 2012.

[4] Winnick, J.P., "The performance of visually impaired youngsters in physical education activities: Implications for mainstreaming" Adapted Physical Activity Quarterly, 2(4), 292-299, 1985.

[5] Zuzanna Maćkowiak, Wieslaw Osiński,"The Effect of Sensorimotor Training on the Postural Stability of Visually Impaired Women Over 50 Years of Age" Journal of Women & Aging, 27(1), 68-80, 2015. DOI: 10.1080/08952841.2014.928140

[6] Wiskochil, B. et al. "The effects of trained peer tutors on the physical education of children who are visually impaired" Journal of Visual Impairment & Blindness, 101(6), 339-350, 2007.

[7] Malik, S. et al., "Orientation and mobility training in special education curriculum for social adjustment problems of visually impaired children in Pakistan" International Journal of Instruction, 11(2), 185-202, 2018.

[8] Bischof, A. et al. "Exploring the playfulness of tools for co-designing smart connected devices: a case study with blind and visually impaired students" Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion,. ACM, 2016.

[9] Lefeuvre, K. et al. "Loaded dice: exploring the design space of connected devices with blind and visually impaired people" Proceedings of the 9th Nordic Conference on Human-Computer Interaction, ACM 2016.

[10] Ando, B. et al. "A haptic solution to assist visually impaired in mobility tasks" IEEE Transactions on Human-Machine Systems, 45(5), 641-646, 2015.

[11] Kraus, S. "Astronomy for the blind and visually impaired" Bringing the Universe to the World, 36, 2016.

[12] Gibby, D. G. U.S. Patent No. 9,133,759. Washington, DC: U.S. Patent and Trademark Office. 2015.

[13] Bagwan, F. et al. "Tangible user interface based descriptor for appliances" National Conference NCPCI, 2016.

[14] Verma, N., Gupta K., & Mahapatra S. "Implementation of solid state relays for power system protection" International Journal of Scientific & Technology Research 4(6), 65-70, 2015.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 309-319 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Sustainable Multi-layered Open Data Processing Model for Agriculture: IoT Based Case Study Using Semantic Web for Hazelnut Fields

Şahin Aydin*, Mehmet Nafiz Aydin

Kadir Has University, Department of Management Systems, 34083, Turkey

ARTICLE INFO

Article history:
Received: 01 January, 2020
Accepted: 19 March, 2020
Online: 23 March, 2020

Keywords:
Open Data
Open Data Processing Model for Agriculture
Multi-layered Open Data Processing Model
Semantic Web in Agriculture
IoT in Hazelnut farming

ABSTRACT

In recent years, several projects which are supported by information and communications technologies (ICT) have been developed in the agricultural domain to promote more precise agricultural activities. These projects account for different kinds of key ICT terms such as internet of things (IoT), wireless sensors networks (WSN), cloud computing (CC). These projects are used for different agricultural products; and it is a well-known fact that they can be essential to perform precise agricultural activities for the relevant agricultural products. The implementation of these projects successfully depends on the extent to which various stakeholders provide support by leveraging relevant data, gathered from heterogenous data sources. Agriculture domain has a great number of stakeholders. These stakeholders need sophisticated data and appropriate intelligence to get benefits in order to perform precise agricultural activities. Authors agreed with scholars that "Open Data" idea, which means accessing data published on the web and available in a machine-readable format is an appropriate way to get benefits for precise agriculture by relevant stakeholders. In this paper, authors shall investigate the open data term in an agricultural context, create an open data processing model, and develop an IoT-based solution to gather environmental data from agricultural fields. Authors also show viability of the proposed model by developing an ICT-based solution. Considering the socioeconomic importance of hazelnut for Turkey, the stakeholders of hazelnut domain still have problems such as availability, meaningful, accuracy of the hazelnut related data. Therefore, authors shall focus on hazelnut within the scope of this paper.

1. Introduction

This paper is an extension of work originally presented in the 6th International Conference on Control Engineering & Information Technology (CEIT), in 2018 [1].

Agriculture increases employment that covers more than third of the world population via exportation and domestic income which makes it one of the leading sectors in the world. [2]. Turkey which is the leader country in terms of hazelnut production in the world, has an important role for contributing more information with respect to hazelnut production cycle from this aspect. As well as Turkey is the leader of hazelnut production, it is leading producer of dried apricots, raisins, and dried figs [3]. Since having such a convenient weather quality which only a few countries could have for hazelnut production, Turkey, has 75% of overall production and 70-75% of the exportation. Furthermore, considering that there have been 4 million people correspond with almost 5% of the overall population of Turkey, who are concerned with hazelnut produced over an area of 550-600 thousand hectares, this makes hazelnut quite important in terms of socio-economic aspect [4]. In Turkey, statistics concerning agricultural products, including hazelnut are presented as only hypertext markup language (HTML) pages, spreadsheets, and comma-separated files (CSV) by Turkish Statistical Institute (TUIK). The context of hazelnut data presented by TUIK contains just a certain type of statistical information. However, the domain stakeholders such as farmers, experts, researchers, and analysts might generate more detailed and sophisticated data regarding hazelnut so that gaining knowledge from it must be straightforward and obvious. In order to achieve that, the data must include basic parameters used for the general management of the accession and environmental and field-specific parameters. This kind of data, freely available to use and

*Şahin AYDIN, Email: aydinoglu.sahin@gmail.com

republish and in machine-readable format, might give a chance to stakeholders for removing the obstacles of implementation of sustainable hazelnut production. Examples of issues regarding hazelnut, which might be commonly encountered by domain stakeholders, might be indicated as follows: limited information or uncertainty about supply and demand side. In addition, there are several problems regarding hazelnut production such as aging hazelnut gardens, land element and position, soil texture, blights, fertilizing, treatment, less pollination, less pruning, agricultural pests, price fluctuation, increase of hazelnut production areas uncoordinatedly, and lack of persistent exportation policies [5]. Furthermore, there exist positive or negative correlations between environmental factors and hazelnut production. For instance, a sudden drop in the temperature negatively affects hazelnut production. It is not generally possible to access this type of data for such a particular agricultural product. However, it might be a solution to leverage the reference body of knowledge and analyze the reports of scientific experiments by accessing this data. The data obtained from scientific experiments are not usually allowed for using by any stakeholders out of the relevant communities [6]. However, considering the presence of barriers for accessing such kind of data, most of organizations belonging to government or private sector are in charge of removing them and making the data freely accessible. Nevertheless, aforementioned permission barriers should be removed for hazelnut stakeholders to meet requirement accessing, reusing, and publishing sophisticated data regarding hazelnut as well.

In this paper, we shall look in detail into processing open data gathered from different data sources and a designed model, considering very exclusive layers from stakeholders' perspectives.

This paper begins with an introduction section in which the research problem is explained. Second section is the research background section that mentions some key terms regarding the model's layers. Then, open data processing model is disclosed in the model, followed by the findings section. Lastly, a case study, which is developed as an exclusive system to demonstrate the model viability, was explained in the implementation section.

2. Research Background

In Turkey, hazelnut production is being dominantly performed in the zone of Black Sea which comprises of two different areas named as First Standard District (FSD) and Second Standard District (SSD) respectively. While SSD is composed of the west and the middle areas of Black Sea zone, FSD is located in the east of the Black Sea zone. Although there are similarities between two areas, it should be noted that considerably differences such as environmental characteristics and climate conditions are exist between them as well. The aforementioned dissimilarities might generate an obvious discrepancy in hazelnut production year by year. When one looks at the statistical data of last thirty years in agricultural production, one sees that on average 520.198 tons hazelnut were produced in Turkey, and such data demonstrates that the position of Turkey as a leader in hazelnut production in the world. Turkey exports significant amount of hazelnut every year. The statistical data regarding agricultural exportation in the year 2019 illustrate that Turkey exported at about 320 thousand tons of hazelnut and it received 2 billion 28 million dollars of revenue via this. It has been identified that there are not adequate researches and studies about how to utilize contemporary ICT for such valuable agricultural product. Therefore, a multi-layered model based on open standards for demonstrating how to process agricultural data is proposed considering the hazelnut which is a dominant agricultural product produced in a wide range area in Turkey in this study. Multi-layered model enables us to distinguish the essential layers such as data sources, data processing, semantic annotation, data storage, services, applications and users functions physically and logically. In addition, it enables interoperability for each layer to accomplish performing sustainable agricultural activities. There are a number of important benefits of the multi-layered approach. For instance, the multi-layered approach is secure, and its management is easy. It is also scalable, flexible and provides more efficient development abilities. Adding new features to the developed system is easy as well. That is why, a multi-layered model is presented and detailed within this study.

Although extensive research has been carried out on agricultural ICT technologies, not many studies exist which adequately focus on hazelnut. Furthermore, there is not a model which meets the requirements of hazelnut agricultural activities as well. So, a multi-layered open data processing model which aims to conduct sustainable agricultural activities of hazelnut is proposed as a part of this study.

The model proposed is essentially based on two pillars: the very idea of "Open Data", and "Semantic Web" approach to data modeling. Authors shall first provide a basic understanding of these pillars so that conceptual foundation of the proposed model is established. Publishing the data in machine-readable format using favorable open standards and open formats by enabling stakeholders to access, read, reuse, and share it with no license limit or permission barriers might be described as "Open Data" [7]. Furthermore, ICT's role is to facilitate distribution of data to users from various stakeholders in the relevant domain. Eliminating heterogeneity of data sources in the agricultural domain might be possible by using the strength of semantic web technologies. It is necessary to publish data using open standards and in machine-readable format to meet requirements of providing processable data by software agents. Semantic Web (SW) aims transforming data into suitable format which might be processed by software agents and human via software tools [8]. As well as SW eases retrieving the data by human or software agents via making interrelations between the objects. The linked data, which might be usable in varied formats on the web and published under open license, must be in structured and machine-readable format [9]. Resource Description Framework (RDF) is used to display information on the web as guided and identified data. It usually serves data including, but not limited to, personal information, social networks, metadata about digital artifacts. Furthermore, it contributes as a way of combination over diverse sources of information. The structure of SPARQL which is being used for querying RDF data, is described semantically considering this specification [10].

Utilizing semantic web technologies is quite efficient way to present domain-specific data on the web. It might ensure removing difficulties in terms of presenting data in appropriate formats, which can be processed by both software agents and humans. RDF, which is widely used and known as common standard for sharing data semantically on the web, has strong characteristics to ease merging data obtained from heterogeneous sources. Merging data obtained from varied sources might be performed using ontologies. RDF, which plays a significant role in SW's world in terms of

presenting data, is main pillar in terms of generating ontology regarding varied domains [11]. RDF also enables improving schemas in time with no requirement for changing by data consumers [12].

3. Model and Findings

As mentioned in the previous sections, authors created a multi-layered open data processing model, and focused on the hazelnut agricultural product. Even though this model focuses on a sustainable agricultural production lifecycle of hazelnut in this study, it is convenient for all agricultural products as well. This model is created by considering data sources, data processing, semantic annotations, data storage, services, applications, and users. The proposed open data model which demonstrates how to process data obtained from heterogeneous agricultural data sources and designed as multi-layered includes varied layers such as types of data sources (IoT devices, statistical data from government, market data, farmers, and other data sources); storing and processing raw data; semantic annotation layer (agricultural product ontology, other ontologies and data interchange engine); data storage layer (graph databases, relational databases, cloud databases, RDF and XML file-based storage options and database service); services layer (XML web services, REST web services, web APIs, mobile services, analysis services, reporting services and SPARQL query services); applications layer (unified web based data platform based on open formats, mobile applications and desktop applications); and lastly, end users layer. The overall model and its each layer, proposed in this study, are being demonstrated in Figure 1.

The data are generally gathered from heterogeneous sources such as farmers, sensors, government statistical data, and market data in the agriculture domain. There is no doubt that the most important one of these heterogeneous data sources is “farmers”. As a significant data source in terms of the model, farmers, play a vital role for performing agricultural production life cycle. Considering the hazelnut production cycle, there is much varied information regarding hazelnut might be collected by consulting of farmers.

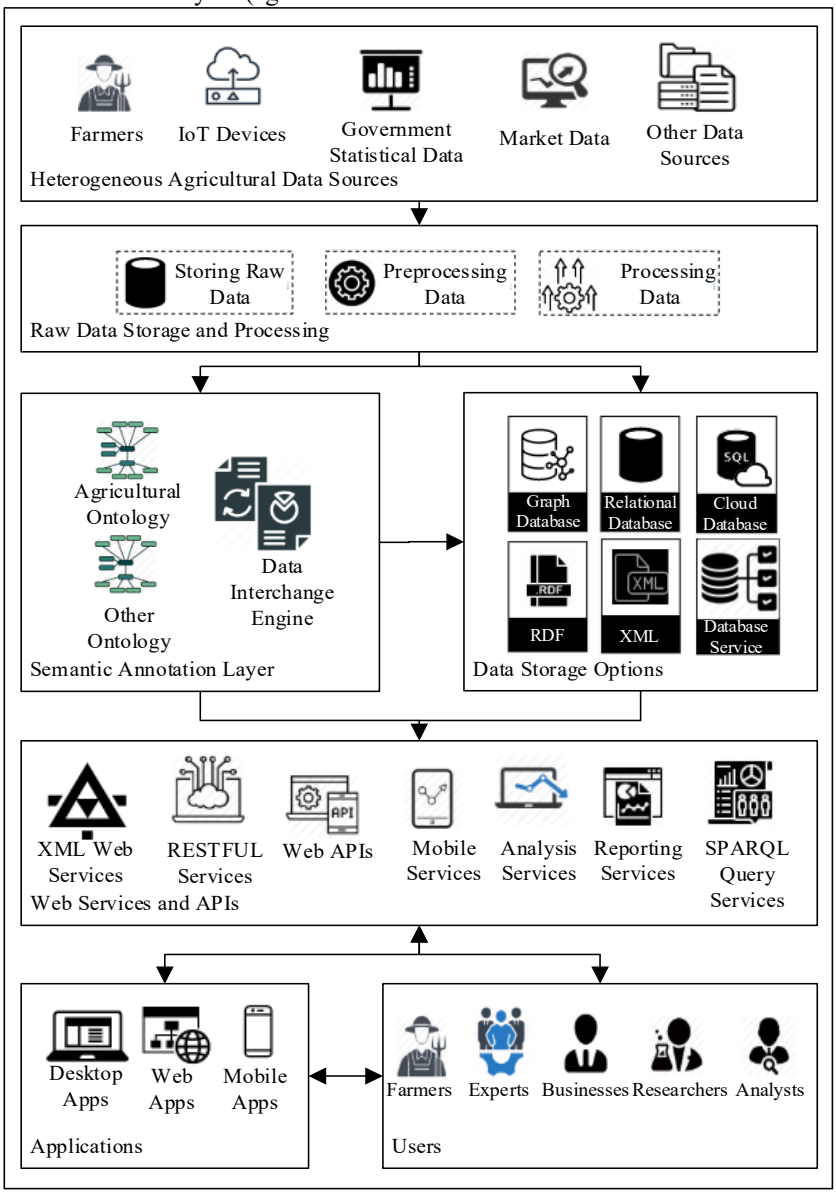


Figure 1: Overall Architecture of Open Data Model

For instance, age productivity and morbidity of trees, observation for soil fertility, the number of trees in orchards, location of orchards (southern slopes or northern slopes), pruning hazelnut trees, fertilization, agricultural spraying, cropping system, propagation method, irrigation method, overall vegetation surrounding the collecting site, stoniness, rockiness etc.

Today, IoT devices have been frequently used for varied domains such as smart cities, industries, education, health, and agriculture. These devices generally consist of a wide range of sensors which are used to collect specific data pertaining to environmental characteristics of site. Collecting and processing data with respect to these characteristics facilitates performing precise agricultural activities. Sensors might provide sensitive measurement values concerning slope, weather temperature, velocity of wind, rainfall, soil moisture, soil Ph, frost (date of most recent frost, minimum temperature, duration of temperature below 0°C), relative humidity (diurnal and seasonal range), light intensity, leaf anatomy, etc.

Agricultural production stats, which are a sort of official statistics data, are recorded by government agencies. The open data model takes into consideration the importance of statistical data concerning agricultural production so, it adopts them as an element of overall architecture. The statistical records regarding many varied domains in Turkey are compiled, evaluated, and published by TUIK [13]. The data with respect to agriculture domain are generally published by TUIK as spreadsheet, comma-separated values, and HTML files, which might be accessed through TUIK web platform.

Marketing operations in agriculture domain include several procedures such as production plan, burgeoning, reaping, classifying, packaging, shipping, managing storehouse, transforming agricultural products into food, delivery, promote, and sale, required for movement of a product from farmers to buyers [14]. It is a very crucial cycle for any agricultural product, so the proposed model contains it as a data source element. The experts, researchers, and analysts who are interested in relevant agricultural products might contribute to the system by loading some exclusive data. So, these kinds of data sources are defined as other data sources in the open data model.

According to NIST (National Institute of Standards and Technology), "cloud computing (CC) is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources such as networks, servers, storage, applications, and services, that can be rapidly provisioned and released with minimal management effort or service provider interaction" [15]. Considering the strength and robust characteristics of CC, it has been recommended that data storages, services and applications of the proposed model should serve on the cloud. Raw Data Storage and Processing Layer (RDSandPL) comprises component of raw data storing (SRD) , component of data preprocessing (PPD), and component of data processing (PD). Data obtained from heterogeneous agricultural data sources are stored into varied data storage options such as relational databases, graph databases, cloud databases, RDF files, and XML files via SRD. It is worth bearing in mind that any stakeholder might need to handle data obtained from heterogeneous sources in raw form. That's why, it is significant storing data in raw format using recommended data storage options according to open data model point of view.

Considering the data gathered from heterogeneous agricultural data sources might be incomplete, incorrect, or irrelevant, it is important to emphasize the requirement of a mechanism to detect and correct them. According to open data model, this mechanism is being corresponded with PPD component. Preprocessing the data ensures removing incompleteness and inconsistency with respect to the raw data and transforming such data into meaningful format to analyze effectively [16]. As well as, PPD decomposes raw data, it aims to apply exclusive techniques concerning preprocessing to such data. Applying preprocessing techniques such as data integration, data enhancements, data enrichment, data transformation, data reduction, data discretization and data cleansing boosts quality of the data [17]. Some of the traditional data preprocessing methods such as K-neighbors Classifier, Logistic Regression, Gaussian Naïve Bayes, Decision Tree, and Support Vector Machine might be used for sensor data, but it should be noted that the major purpose of development of these methods are not to apply them to sensor data [18]. This component is also responsible for semantic annotation of streaming data obtained from sensors, using existing or new ontologies. By this means, semantic annotation of streaming sensor data facilitates data integration processes between varied data sources and open data platforms. In addition, this component solves noisy and redundant problems of sensor data streams, solves missing value problems, detects anomalies related to data, and manages data heterogeneity using the strength of semantic web technologies [19].

The aim of carrying out the preprocessing techniques and methods to the sensor stream data is to build a robust technical infrastructure for data processing component of the proposed model. Data processing means that acquiring meaningful information from preprocessed raw data. Acquiring meaningful information enables users to make estimations for the future via implementing varied analyze techniques and allows creating analytical solutions. So, it is important to bear in mind that having such a component is essential in terms of open data model. The design principles proposed by Begoli and Horey, which allows using varied analytical techniques and substantially interested in maximizing the controllable factors, provide easily exploring and analyzing the data by researchers [20]. The practical analytical solutions are ensured corresponding with these design principles for stakeholders by component of data processing of open data model.

When implementation of relevant methods and techniques is completed by the components of RDSandPL, it is required transforming data into RDF datasets. The data interchange component of Semantic Annotation Layer (SAL) is in charge of transforming data into RDF datasets. The proposed model recommends using agricultural product ontologies to transform data into RDF datasets.

Ontologies distinctly specify conceptualization by means of constructing abstract and simplified view of a particular domain which requires to be represented [21]. Combining a number of objects and concepts, which are related to knowledge that is supposed to be expressed and making interrelations between them might be described as conceptualization [22]. The aim of generating ontologies is not merely to identify the conceptualization or point out the widespread savvy of a particular domain. In addition, they are utilized as conceptual models which provide common understanding to facilitate communicating

between varied types of software application systems and among people [23]. Nowadays, there exist several ontologies which are still being in use, concerning the agriculture domain. AGROVOC, one of the commonly known agricultural ontologies, was developed by Food and Agriculture Organization (FAO) of the United Nations in the early of the 1980s. Furthermore, it is the most popular and well-known agricultural thesaurus all over the world. AGROVOC, which is formed as linked dataset, is based on SKOS-XL (Simple Knowledge Organization System eXtension for Labels) concept scheme, and it comprised of more than thirty-five concepts which are defined in up to twenty-nine different languages. [24]. Institute for Learning & Research Technology created SKOS(Simple Knowledge Organization System) with the aim of providing linked and distributed knowledge organization systems such as controlled vocabularies, thesauri, taxonomies, and folksonomies [25]. SKOS-XL, as an extension version of SKOS, facilitates identifying lexical entities and making interrelationships among them, with the additional support [26]. The focus of this study is on the development of open data processing model. Therefore, authors intend to develop a comprehensive ontology using descriptors of hazelnut as future work [27]. Both RDF (Resource Description Framework) and OWL (Web Ontology Language) have vital role in the world of semantic web and are popular technologies which might be utilized to accomplish the aforementioned objective. RDF and OWL are the most known standards of semantic web technologies for describing data on the web. OWL is developed based on RDF and extends it including exclusive components such as axioms, classes, object properties, data properties, datatypes, individuals etc. [28]. OWL is the common, well-known, and standard language to create ontologies. Moreover, it has three sublanguages to meet different requirements such as OWL Full, OWL DL, and OWL Lite.

Data gathered from different data sources and processed by RDSandPL might be stored in graph databases, relational databases, cloud databases, XML files, and RDF files, which are recommended as data storage options in the open data model. The proposed open data model is noted the importance of converting raw data into RDF datasets. Therefore, transforming data gathered from varied sources into triple store which means subject, predicate, and object, has been built as a component of semantic annotation layer. Due to RDF datasets have robust features in terms of storage and indexing [29], they have been considered as a storage option in the proposed open data model. Relational databases component of data storage layer (DSL) is in charge of storing relational data. Graph databases component of DSL is in charge of storing data as graphs. Cloud databases might store data as both graphs and relational. Database services, which is a component of DSL, provide a secure connection, session management, user mapping, user authentication, and authorization between components of services layer (SL).

Another essential layer of proposed model is the services layer, which consists of seven different types of services: XML web services, REST web services, web APIs, mobile services, analysis services, reporting services, and SPARQL query services. According to W3Schools, "web services are defined as self-contained and self-describing application components that can be discovered using UDDI, be used by other applications, and they communicate using open protocols" [30]. Web Services and APIs Layer (WSandAL), which is in charge of providing interoperability between mobile, desktop, and embedded software applications and databases, is a significant part of the proposed model. While providing interoperability by WSandAL, the RESTful Services component of this layer plays an important role. Software agents, tools and applications, which are being developed using different or similar development technologies, utilize agricultural data by consuming RESTful Services. Another important feature of the proposed model is mobility. To meet this requirement, it strongly recommends developing mobile applications. It is, therefore, needed to generate mobile services to develop an efficient infrastructure for mobile devices. Mobile services, which provide connection to different kinds of mobile devices must take part in the proposed model. Stakeholders might be limited by lack of knowledge to apply data analysis methods to agricultural data. The proposed model recommends building analysis services to solve this problem.

Designing and publishing detailed reports are required to develop usable software systems which aims to process data. Reporting Services component of open data model aims to meet these requirements. It enables domain stakeholders create dynamic, well-designed and detailed reports. As mentioned in previous section of this study SPARQL is the language to query RDF datasets which are stored in files or graph databases. In the proposed open data model, creating queries using this particular language and compiling them are provided and facilitated by service component named SPARQL Query Service.

Applications Layer (AL) allows end users to access agricultural data obtained from heterogeneous data sources. AL consists of web, mobile and desktop applications which enable us to import, export, report, analyze, and query data.

There are a number of varied end users belonging to proposed open data model, called domain stakeholders such as farmers, researchers, businesses, analysts, and domain experts. Domain stakeholders are provided to utilize data which are gathered from heterogeneous agricultural data sources and might be consumed through the WSandAL and processed by applications layer. Considering the importance of accessing, reusing, and processing such data, it allows domain stakeholders of hazelnut agricultural product construct common knowledge with respect to performing agricultural activities in an efficient way. Every year, several exploratory investigations are performed to estimate total amount of hazelnut production by different institutions of Turkey. However, these current exploratory investigations to find out the hazelnut production of a further year are limited because of unforeseen circumstances with respect to hazelnut harvest. Unfortunately, some discrepancies are seen among the estimations which are carried out by different government agencies every year. Due to the hazelnut price per kilogram is being affected from these discrepancies negatively, rapid price fluctuations are being seen for hazelnut in Turkey. When one takes into consideration the problems for hazelnut market, each user in other words domain stakeholder, play a vital role in terms of the proposed open data model.

The proposed open data model reveals that how each of stakeholders might exploit the processed and published agricultural data for removing the obstacles and problems concerning any particular agricultural product. The proposed open data model by its data processing perspective might enable farmers who have a key role in agriculture domain plan their production cycle effectively. The domain experts and researchers might take preventive measures to prevent further spread of the agricultural

diseases by producing the meaningful information using the applications and services developed based on the proposed open data model. Varied kinds of statistical analysis with respect to market, investment, and employment might be undertaken by analysts utilizing the conferred data by proposed open data model. Statistical analysis presented by analysts might be used to support making effective decisions concerning market share of agricultural businesses for the future.

4. Implementation

The proposed multi-layered open data processing model is introduced within the previous sections. In this part of the study, a prototype of the electronic system is developed by establishing a wireless sensor network (WSN), which consists of varied sensors and other electronic devices in order to test the viability of the proposed model. This implementation is only focused on streaming data from IoT devices to the databases through web services. Other aspects of the proposed model shall be elaborated in different studies in future studies. There are three sensor nodes named router 1, router 2, and router 3 in the developed system, respectively. These nodes transmit measured data to the coordinator. This system structure is an implementation of ordinary WSN. Figure 2 presents the architecture of the developed wireless sensor network.

Weather temperature, weather humidity, weather pressure, carbon monoxide, nitrogen, oxygen, and ultraviolet detection sensors have been plugged on router 1; digital light intensity, soil moisture, and precipitation detection sensors have been plugged on router 2; and accelerometer and gyroscope sensor have been plugged on router 3. Table 1 illustrates which sensor measures what type of measurements, and which sensor is plugged on which router. While some of these sensors measure different types of environmental events or changes at the same time, some of them measure only one event or change. Weather pressure, temperature, humidity changes have been detected by only one sensor named BME280. Likewise, rates of carbon monoxide, hydrocarbons and oxidizing gases such as NO_2 have been measured by only one sensor named MICS-4514. Arduino Uno microcontroller boards based on ATmega328P have been used to plug the sensors and measure their values in the system. XBee is one of the most popular modules which allows devices connect to each other through wireless. So, XBee Pro S2C modules have been seen an appropriate embedded solution for established wireless sensor network to transmit data from routers to the coordinator. The coordinator consists of a Raspberry Pi 3 small single-board computer, an Arduino Uno, and an XBee Pro S2C module. Raspberry Pi has been included in WSN to run a universal windows application, which is in charge of reading, visualizing, and storing real-time data stream obtained from sensors. This application communicates with the coordinator device to read the data stream through serial port. The real-time data stream, which is visualized in favorable visualization objects by universal application, are stored in databases through web services. Figure 3 illustrates how sensor measurements are visualized by universal windows application on Raspberry Pi.

ZigBee is an international open standard and communication protocol, which depends on IEEE 802.15.4 and ensures low-power, low-cost, low-data-rate, and wireless mesh networking. [31]. Zigbee has three kinds of node types, such as coordinator, router, and end device. However, the developed WSN does not include any end device. Each of the XBee modules belonging to WSN were configured to API operating mode, which provides data transmission to multiple destinations with no requirement, using command mode. Receiving success and failure status of each transmitted RF packet is facilitated in this mode. In addition, the operation of identifying the source address of each received packet is readily provided by this mode. The routers transmit data, which is gathered via sensors, in an exclusive format named frame, represented as a ZigBee packet to the coordinator. The structure of the frame data; in other words, a ZigBee packet, which is used by the routers to transmit sensor data stream towards the coordinator, is demonstrated in Figure 4. The structure of frame data differs in compliance with the objectives of the API frame. The frame data structure of the established wireless sensor network, represented in Figure 4, consists of the fields, such as start delimiter, length, frame type, frame id, 64-bit destination address, options, RF data, and checksum.

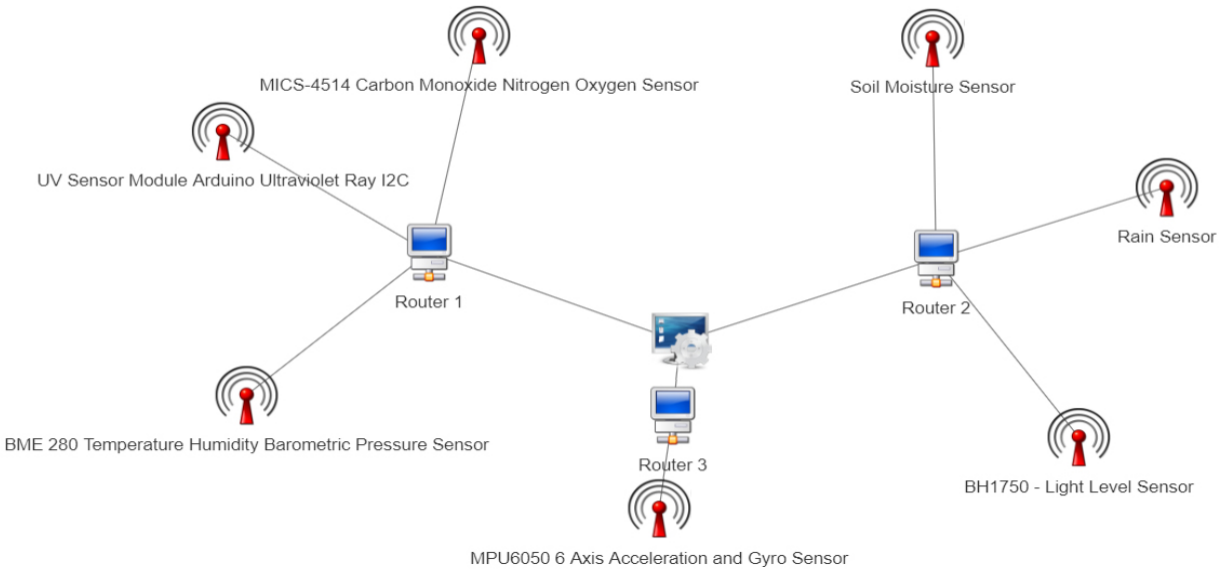
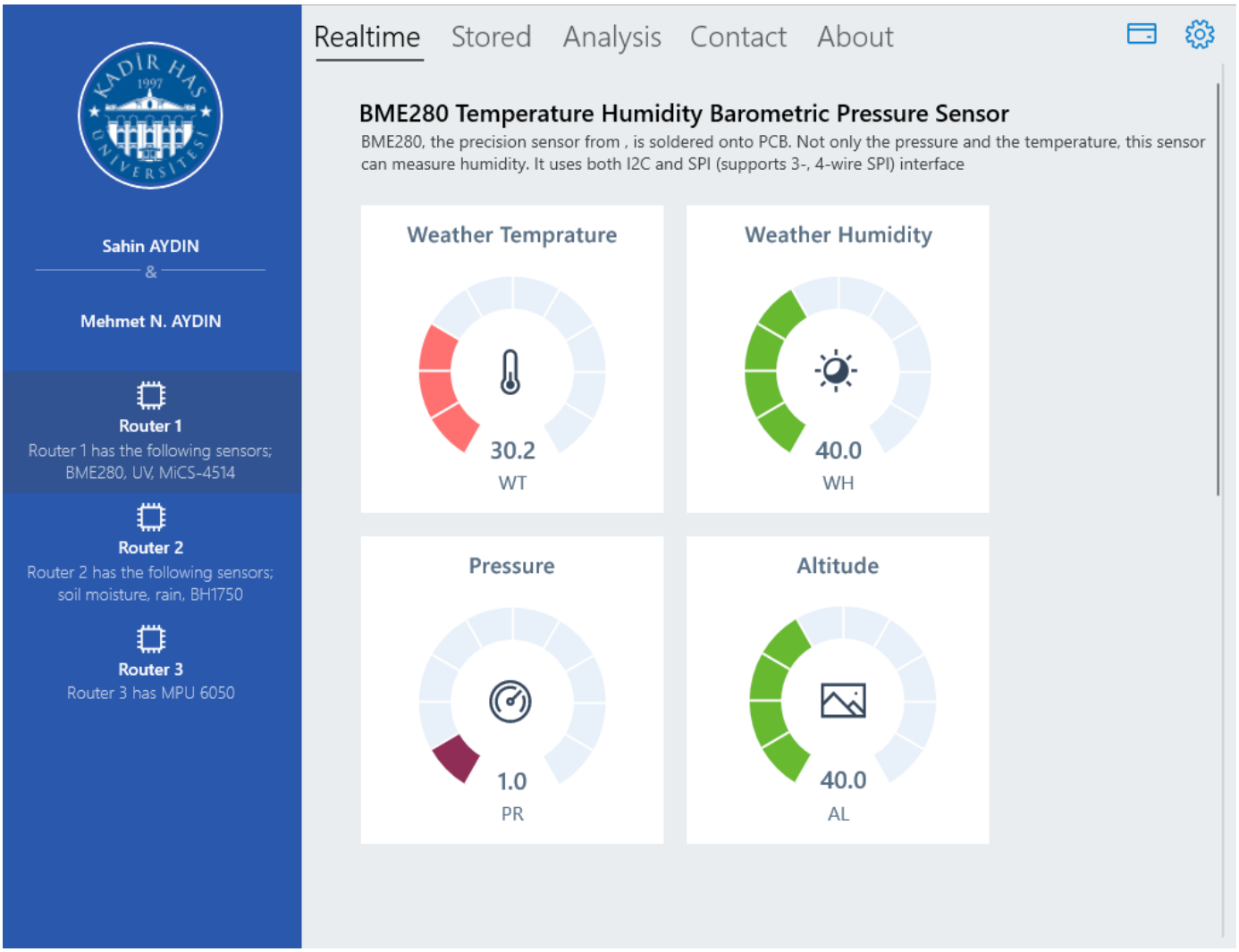


Figure 2: Developed WSN Architecture

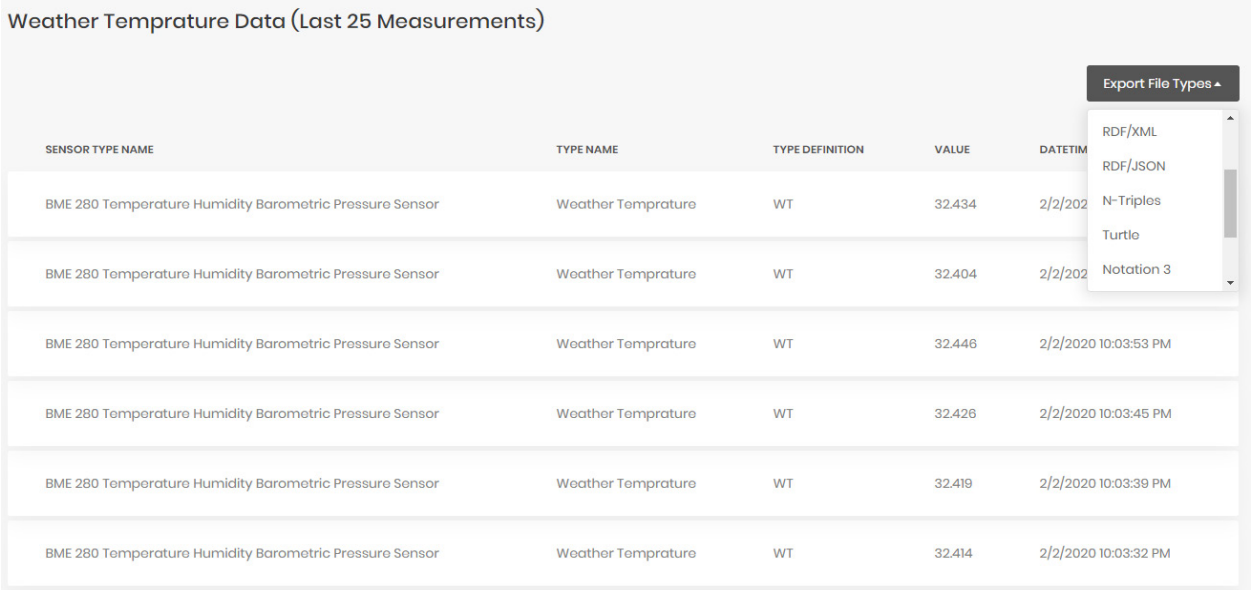
Table 1: Sensors plugged on routers

Router Name	Sensor Name	Measurements
Router 1	BME280 Temperature Humidity Barometric Pressure Sensor	Temperature
		Humidity
		Pressure
	GUVA-S12SD UV Sensor Module Arduino Ultraviolet Ray I2C	UV radiation in sunlight
	MICS-4514 Sensor	CO concentration
		Hydrocarbons
		Oxidizing gases
Router 2	Soil Moisture Sensor	Soil Moisture
	BH1750 - Light Level Sensor	Light Intensity
	Rain Sensor	Rain detection and rainfall intensity measurement
Router 3	MPU6050 (Gyroscope + Accelerometer + Temperature) Sensor Module	Accelerometer Gyroscope Temperature



7E	00 1F	00	01	00 00 00 00 00 00 00 00	00	52 49 44 3A 31 3B 53 49 44 3A 31 3B 57 54 3A 33 32 2E 32 35	14
Start Delimiter	Length	Frame Type	Frame Id	64-bit destination address	Options	RF Data	Checksum

Figure 4: Frame Data Format



The sensor data are transmitted from the routers to the coordinator using key-value pair identifiers. For instance, the frame data, represented in Figure 4, refers to "RID:1;SID:1;WT:32.25". The key-value pairs of this frame are RID:1, SID:1, and WT:32.25, respectively. Likewise, the values are 1, 1, and 32.25 as well. The keys, RID, SID, and WT stand for router id, sensor id, and temperature, respectively.

When the data are transmitted to the coordinator and transformed into a meaningful format by universal application, storage operations are performed by DPL. Data are stored in two different types of database management systems, such as Microsoft MS SQL Server and Neo4j graph database system, which are a component of data storage layer. Neo4j, which is a graph database management system, increases the system performance, provides flexibility and agility. Figure 6 presents the data stored in Neo4j. RDF file-based data storage option is used to store and export sensor data for this case study. Table 3, Table 4, Table 5 represent RDF/XML, Notation 3, and N-Triples datasets, respectively. These datasets can be created using the sensor listing UI of the open data platform represented in Figure 5.

The data gathered from sensors and stored on databases must be published in a suitable format for applications running on different platforms. Therefore, REST web services were developed using Windows Communication Foundation (WCF), which is a framework for building service-oriented applications. Furthermore, web services for Windows platforms were developed in the same way. Table 2 illustrates the formatted data in JSON (JavaScript Object Notation) format. Developed web services are accessible to stakeholders.

Table 2: Example of JSON data

```
[{
        "Id": 2,
        "SensorDataAddedDatetime": "\/Date(1580635128430+0300)\/",
        "SensorDataType": 2,
        "SensorDataValue": "32.45",
        "SensorId": 1000
}]
```

Table 3: Example of RDF/XML dataset

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE rdf:RDF [
        <!ENTITY rdf 'http://www.w3.org/1999/02/22-rdf-syntax-ns#'>
        <!ENTITY rdfs 'http://www.w3.org/2000/01/rdf-schema#'>
        <!ENTITY xsd 'http://www.w3.org/2001/XMLSchema#'>
        <!ENTITY sensorData 'http://www.opendatainagriculture.com/sensors#'>
]>
<rdf:RDF xml:base="http://www.opendatainagriculture.com/sensors" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:xsd="http://www.w3.org/2001/XMLSchema#" xmlns:sensorData="http://www.opendatainagriculture.com/sensors#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:Description rdf:about="http://www.opendatainagriculture.com/sensors/datatype#WeatherTemperature">
    <sensorData:DataTypeId>2</sensorData:DataTypeId>
    <sensorData:Datetime>2/2/2020 10:04:07 PM</sensorData:Datetime>
    <sensorData:SensorTypeId>1000</sensorData:SensorTypeId>
    <sensorData:SensorTypeName>BME 280 Temperature Humidity Barometric Pressure Sensor</sensorData:SensorTypeName>
    <sensorData:TypeDef>WT</sensorData:TypeDef>
    <sensorData:TypeName>Weather Temprature</sensorData:TypeName>
    <sensorData:Value>32.434</sensorData:Value>
  </rdf:Description>
</rdf:RDF>
```

Table 4:Example of Notation 3 dataset

```
@base <http://www.opendatainagriculture.com/sensors>.

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
@prefix sensorData: <http://www.opendatainagriculture.com/sensors#>.

<http://www.opendatainagriculture.com/sensors/datatype#WeatherTemperature> sensorData:DataTypeId "2";
                                        sensorData:Datetime "2/2/2020 10:04:07 PM";
                                        sensorData:SensorTypeId "1000";
                                        sensorData:SensorTypeName "BME 280 Temperature Humidity Barometric Pressure Sensor";
                                        sensorData:TypeDef "WT";
                                        sensorData:TypeName "Weather Temprature";
                                        sensorData:Value "32.434".
```

Table 5: Example of N-Triples dataset

```
<http://www.opendatainagriculture.com/sensors/datatype#WeatherTemperature>
<http://www.opendatainagriculture.com/sensors#DataTypeId> "2”.
<http://www.opendatainagriculture.com/sensors/datatype#WeatherTemperature>
<http://www.opendatainagriculture.com/sensors#SensorTypeId> "1000”.
<http://www.opendatainagriculture.com/sensors/datatype#WeatherTemperature>
<http://www.opendatainagriculture.com/sensors#SensorTypeName> "BME 280 Temperature Humidity Barometric Pressure
Sensor".
<http://www.opendatainagriculture.com/sensors/datatype#WeatherTemperature>
<http://www.opendatainagriculture.com/sensors#TypeName> "Weather Temperature”.
<http://www.opendatainagriculture.com/sensors/datatype#WeatherTemperature>
<http://www.opendatainagriculture.com/sensors#TypeDef> "WT”.
<http://www.opendatainagriculture.com/sensors/datatype#WeatherTemperature>
<http://www.opendatainagriculture.com/sensors#Value> "32.434”.
<http://www.opendatainagriculture.com/sensors/datatype#WeatherTemperature>
<http://www.opendatainagriculture.com/sensors#Datetime> "2/2/2020 10:04:07 PM”.
```

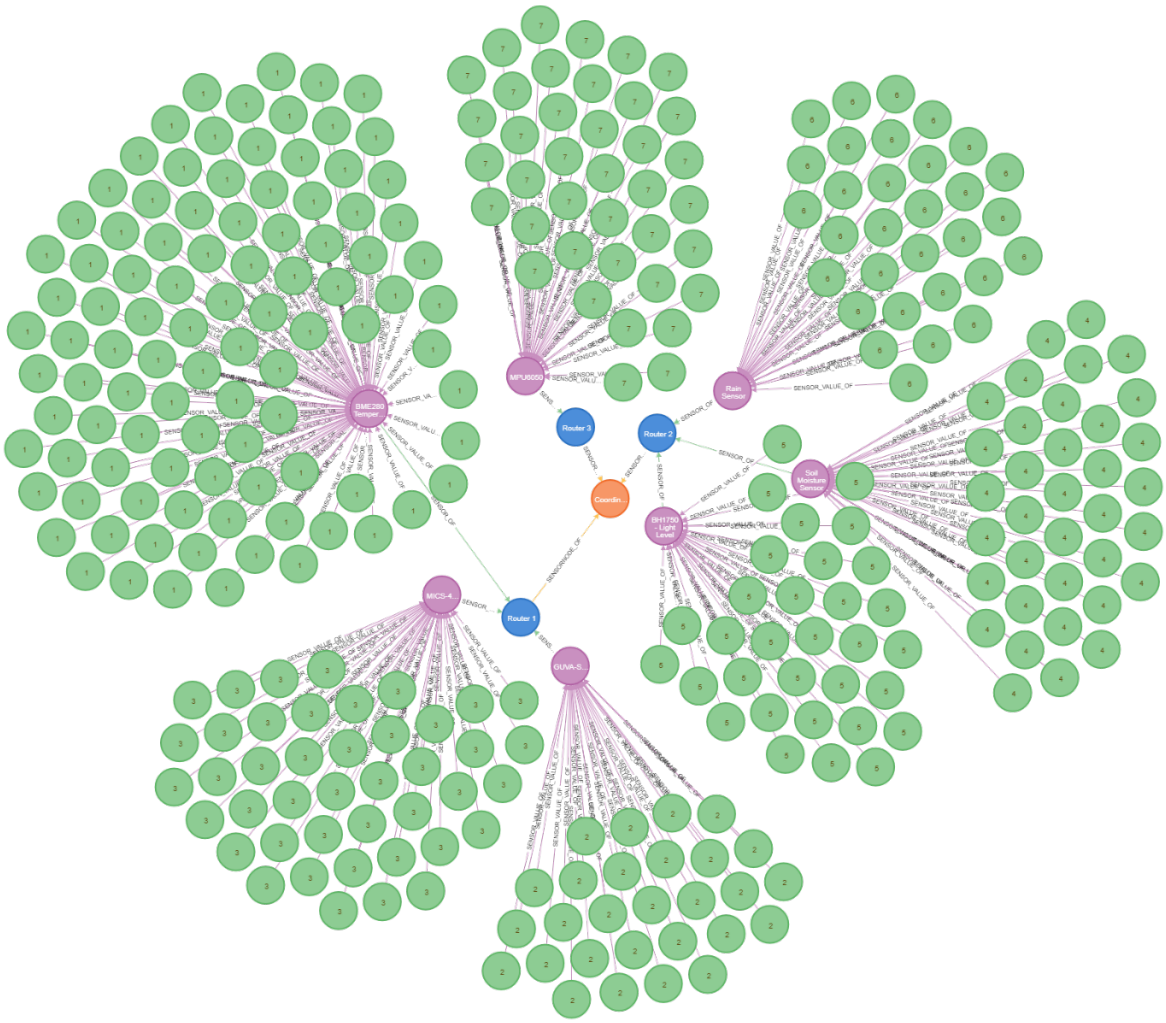


Figure 6: Storing sensor data in Neo4j Graph Database

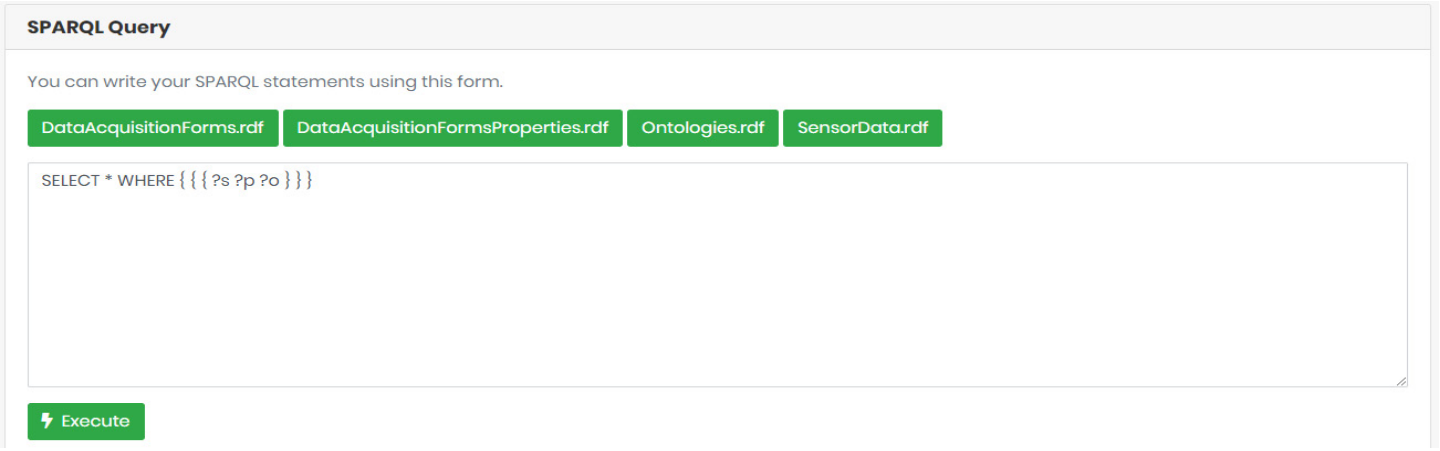


Figure 7: SPARQL query service

Created RDF datasets might be manipulated using SPARQL query language through SPARQL query service, which is represented in Figure 7. This service filters files in compliance with the ".RDF" extension and creates buttons for each of them. Stakeholders should click to load the relevant dataset on SPARQL query service engine. After the loading process is completed, it should be written to execute the query into the text area.

5. Conclusion and Future Work

In this paper, authors proposed a multi-layered open data processing system that consists of different layers and components. All layers and its components were explained; and an implementation regarding sensors was developed and summarized in the implementation section. Authors examined only one aspect

in this implementation. However, there are four other aspects in terms of heterogeneous data sources. These other aspects are worth discussing and examining in the future. Also, authors tried to justify the suitability of the data life cycle in terms of the proposed multi-layered model by implementing the wireless sensor network system. Authors feel that this study enhances academic understanding of creating a multi-layered open data processing model. However, the life cycle of other data from other data sources should be examined and justified for viability. In addition, authors mentioned creating hazelnut ontology, but did not its development and description. As a future work, an ontology for hazelnut shall be created and adapted to the model. In addition, authors are planning to develop an open data platform based on the proposed model. Stakeholders will be able to create data acquisition forms using agricultural trait dictionaries to gather structured data regarding the relevant agricultural product without requiring strong technical skills. The platform will allow obtaining environmental data from heterogenous sensors and will be able to map these sensor measurements with relevant classes of agricultural trait dictionaries. By this means, domain-specific data will be able to be published in varied open formats using the developed platform. A number of restrictions of this study and areas for future research should be mentioned; and the following ones might be given as examples: difficulty of gathering data from farmers (lack of knowledge of technology usage), defining the boundaries of gathering market data, determining the potential data sources, which is defined as "*Other Data Sources*" in heterogeneous data sources layer, and spreading the usage of applications mentioned in "*Applications Layer*".

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This project is partly supported by Scientific Research Fund (2018-BAP-14) of Kadir Has University.

End notes

Web-based Open Data Platform can be accessed using the following URL: http://www.opendatainagriculture.com/

References

[1]. Ş. Aydın, U. Ünal and M. N. Aydın, "Open Data in Agriculture: Sustainable Model Development for Hazelnut farms using semantics," 2018 6th International Conference on Control Engineering & Information Technology (CEIT), Istanbul, Turkey, 2018, pp. 1-6.doi: 10.1109/CEIT.2018.8751875

[2]. Melissa J. Perry, Agricultural Health and Safety, Editor(s): Stella R. Quah, International Encyclopedia of Public Health (Second Edition), Academic Press, 2017, Pages 38-44, ISBN 9780128037089

[3]. Agriculture and Food. [online] Available at: http://www.invest.gov.tr/en-US/sectors/Pages/Agriculture.aspx [Accessed 21.07.2018].

[4]. Turkey's Hazelnut. [online] Available at: http://www.ftg.org.tr/en/turkish-hazelnutturkeys- hazelnut.html [Accessed 21.07.2018].

[5]. Fındık Veriminde ve Üretiminde Temel Sorunlar. [online] Available at: http://www.findikihracati.com/findik/findik-veriminde-ve-uretiminde-temel-sorunlar.html [Accessed 21.07.2018].

[6]. Peter Murray-Rust, Open Data in Science, Serials Review, 34:1, 52-64, 2008, https://doi.org/10.1080/00987913.2008.10765152.

[7]. Carolan, L., Smithy, F., Protonotarios, V., Schaap, B., Broad, E., Hardinges, J., Gerry, W., 2015, How can we improve agriculture, food and nutrition with open data? , Open Data Institute.

[8]. I. Szilagyi and P. Wira, "Ontologies and Semantic Web for the Internet of Things - a survey," IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, Florence, 2016, pp. 6949-6954.https://doi.org/10.1109/IECON.2016.7793744.

[9]. Berners-Lee, T., 2006, Linked Data, https://www.w3.org/DesignIssues/LinkedData.html [Accessed 23.07.2018].

[10]. SPARQL 1.1 Query Language. [online] Available at: https://www.w3.org/TR/2013/REC-sparql11-query-20130321/ [Accessed 23.07.2018].

[11]. Lata, Swaran & Sinha, Bhaskar & Kumar, Ela & Chandra, Somnath & Arora, Raghu, Semantic Web Query on e-Governance Data and Designing Ontology for Agriculture Domain. International Journal of Web & Semantic Technology. 4., 2013 http://doi.org/10.5121/ijwest.2013.4307.

[12]. Resource Description Framework. [online] Available at: https://www.w3.org/RDF/ [Accessed 23.07.2018]

[13]. Available at: http://www.turkstat.gov.tr/UstMenu.do?metod=gorevYetki [Accessed 24.07.2018].

[14]. Agricultural Marketing. [online] Available at: https://en.wikipedia.org/wiki/Agricultural_marketing [Accessed 24.07.2018]

[15]. Mell, P., G., Tim, 2018, National Institute of Standards and Technology, Information Technology Laboratory – a nonregulatory federal agency within the U.S. Department of Commerce – http://csrc.nist.gov/groups/SNS/cloud-computing.

[16]. Aisha Siddiqa, Ibrahim Abaker Targio Hashem, Ibrar Yaqoob, Mohsen Marjani, Shahabuddin Shamshirband, Abdullah Gani, Fariza Nasaruddin, A survey of big data management: Taxonomy and state-of-the-art, Journal of Network and Computer Applications, Volume 71, 2016, Pages 151-166, ISSN 1084-8045, https://doi.org/10.1016/j.jnca.2016.04.008.

[17]. I. Taleb, R. Dssouli and M. A. Serhani, "Big Data Pre-processing: A Quality Framework," 2015 IEEE International Congress on Big Data, New York, NY, 2015, pp. 191-198. https://doi.org/10.1109/BigDataCongress.2015.35.

[18]. Zhong Y, Fong S, Hu S, Wong R, Lin W. A Novel Sensor Data Pre-Processing Methodology for the Internet of Things Using Anomaly Detection and Transfer-By-Subspace-Similarity Transformation. Sensors (Basel). 2019;19(20):4536. Published 2019 Oct 18. https://doi.org/10.3390/s19204536.

[19]. Jason J. Jung, Semantic preprocessing for mining sensor streams from heterogeneous environments, Expert Systems with Applications, Volume 38, Issue 5, 2011, Pages 6107-6111, ISSN 0957-4174 https://doi.org/10.1016/j.eswa.2010.11.017.

[20]. E. Begoli and J. Horey, "Design Principles for Effective Knowledge Discovery from Big Data," 2012 Joint Working IEEE/IFIP Conference on Software Architecture and European Conference on Software Architecture, Helsinki, 2012, pp. 215-218, https://doi.org/10.1109/WICSA-ECSA.212.32.

[21]. Thomas R. Gruber, Toward principles for the design of ontologies used for knowledge sharing?, International Journal of Human-Computer Studies, Volume 43, Issues 5–6, 1995, Pages 907-928, ISSN 1071-5819, https://doi.org/10.1006/ijhc.1995.1081.

[22]. N. F. Noy and C. D. Hafner, "The State of the Art in Ontology Design: A Survey and Comparative Review", AIMag, vol. 18, no. 3, p. 53, Sep. 1997, https://doi.org/10.1609/aimag.v18i3.1306 .

[23]. Cristani, Matteo & Cuel, Roberta, A Survey on Ontology Creation Methodologies. Int. J. Semantic Web Inf. Syst.. 1.,2005, 49-69. 10.4018/jswis.2005040103.

[24]. LNCS About AGROVOC Page. [online] Available at: http://aims.fao.org/standards/agrovoc/concept-scheme, last [Accessed 2019/01/15].

[25]. Dean Allemang, Jim Hendler, Chapter 10 - SKOS—managing vocabularies with RDFS-Plus, Editor(s): Dean Allemang, Jim Hendler, Semantic Web for the Working Ontologist (Second Edition), Morgan Kaufmann, 2011, Pages 207-219, ISBN 9780123859655, https://doi.org/10.1016/B978-0-12-385965-5.10010-X.

[26]. LNCS SKOS Simple Knowledge Organization System eXtension for Labels (SKOS-XL). [online] Available at: https://www.w3.org/TR/skos-reference/skos-xl.html, [Accessed 2019/01/15].

[27]. Koksal, A.,I., Gunes, N., Descriptors for Hazelnut, Bioversity International and FAO, 2008.

[28]. Pollock, J., T.: Semantic Web for Dummies. 1st edn. Wiley Publishing, Inc., Canada ,2009, ISBN: 978-0-470-39679-7 .

[29]. Luo Y., Picalausa F., Fletcher G.H.L., Hidders J., Vansummeren S., Storing and Indexing Massive RDF Datasets. In: De Virgilio R., Guerra F., Velegrakis Y. (eds) Semantic Search over the Web. Data-Centric Systems and Applications. Springer, Berlin, Heidelberg, 2012, https://doi.org/10.1007/978-3-642-25008-8_2.

[30]. XML Web Services. [online] Available at: https://www.w3schools.com/xml/xml_services.asp [Accessed 25.07.2018].

[31]. DIGI International Inc., XBee/XBee-PRO S2C Zigbee RF Module User Guide, 2018

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 313-320 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Fabrication of Glaze Material from Recycled Bottle Glass and Kaolin

Agus Dwi Anggono*, Elkana Bilak Lopo, Joko Sedyono, Tri Widodo Besar Riyadi

Department of Mechanical Engineering, Universitas Muhammadiyah Surakarta, Jl.Ahmad Yani, PO.BOX 1 Pabelan, Surakarta 57162, Indonesia

ARTICLE INFO

Article history:
Received: 24 October, 2019
Accepted: 22 November, 2019
Online: 12 December, 2019

Keywords:
Glaze
Recycle Glass Bottle
Kaolin
Vickers Hardness

ABSTRACT

This research aimed at developing a method of glaze material making from used glass bottle waste combined with kaolin. Then, the microstructure and hardness level of glaze products were characterized. Furnace heaters were used in research to heat glaze material by varying the holding time of 30, 45 and 60 minutes. The composition of glaze material was varied with the ratio between glass and kaolin of 70:30, 80:20 and 90:10. The results revealed that the process of glaze material making was successfully carried out. The results of the hardness test showed that at the 30 minutes holding time, the hardness was 82.53 VHN. The hardness value increased when the holding time was 45 minutes with an average hardness value of 85.03 VHN. The highest average hardness obtained when the variation of 60 minutes holding time was 87.6 VHN. The difference in hardness value in variation possibly occurred because of the differences in the distribution of amorphous solids, porosity, and pinhole which was influenced by the differences in glass and kaolin composition.

1. Introduction

Ceramic glass products are widely known for their unusual combination properties which have exotic combinations of material properties that produce a variety of high-tech products according to the market needs [1]. S.D Stookey introduced the term "pyroceram" for ceramic glass (glass code 9606) in which pyroceram is an original ceramic glass material developed and licensed by Corning Glass in the 1957s. Prior to the development of pyroceram, Stookey worked in controlled crystallization in glasses that were sensitive to radiation (glass code 8603) and was named photosensitive opal glass. But in 1959, Stookey adopted and used the term "ceramic glass" for two types of material, not depend on the type of nucleation agent used (Cu, Ag, Au or TiO2) [2].

The term ceramic glass indicates a material that contains a significant crystal volume fraction of > 50%. The volume fraction of ceramic glass crystals is significantly lower than that which has been developed over the past 60 years. Therefore, problems that limit the number of crystalline phases or the composition of ceramic glass have been removed. Aluminasilikat system is widely used as an object of research and has been commercialized as ceramic glass materials. The importance of new features in the development of ceramic glass production is the need for evolution control in various types of crystals simultaneously or independently grow in large numbers or on the surface of the sample [3][4].

Ceramic products have been widely applied in engineering, especially in machineries such as cutting tools, nozzles, valves, turbines, and ball bearing [5]. The advantages of ceramics, in general, are a high melting point, high-temperature resistance, friction resistance, corrosion resistance, low heat conductivity, relatively low density and low thermal expansion coefficient [6]. Currently, ceramic materials have been developed into modern products with very varied advantages. Ceramic base material in the market is usually made from clay which has passed the process of molding, heating, and finishing. Clay is one of the basic ingredients of ceramic which is plastic, easy to print, stiff after being dried and glassy after being heated at a certain temperature. The use of clay as a base for ceramics is largely determined by the nature of the mineral of the clay compilers. The nature of clay minerals is determined by several factors such as chemical composition and mineral types, physical and chemical properties, and the structure of clay minerals so that clay quality test as a ceramic raw material can not only be determined by chemical composition testing, but also must be tested through the test for the types, mineral structures and physical properties. The clay is most commonly used as ceramic materials such as red brick, tile,

*Corresponding Author: Agus Dwi Anggono, Email: ada126@ums.ac.id

earthenware, and others. Generally, the tile is made of clay, but it has some weaknesses such as cracking, breaking and selling price is too expensive. So, it is necessary to increase mechanical and physical properties through experimental studies of the material coating.

The data from the Ministry of Environment and the Ministry of Public Works of the Republic of Indonesia in 2017 showed that the total waste in Indonesia reaches 187.2 million tons/ year, where the composition of glass waste reached 17% of 187.2 million tons. Waste is the material that is thrown, will be thrown or no longer used according to its purpose in which a material considered waste if no longer used or it is essentially a kind of waste such as liquid waste. From the data, it can be seen that if the waste cannot be reduced, it will harm the environment and nature. Quoted from *ristekdikti.go.id* about "how to manage urban waste" written by Agus Puji Prasetyono, it is said that the priority of the city or capital to waste becomes "Apiori" in which apriori is existing knowledge before having experience. The term is used to explain that someone can think and have assumptions about everything before having experience and finally draw conclusions. By this, talk about rubbish means to talk about "how". Therefore this becomes a guideline or motivation in selecting the research topic. It should be noted that glass waste takes 10-1000 years for the glass to completely decompose. Based on this, concrete steps must be taken to reduce glass waste so that the amount of glass waste does not increase. Several steps to reduce glass waste are implementing the 3R method involving Reuse, Reduce and Recycle. This method has been carried out by several industries, non-governmental organizations, and individuals who care about the environment and want to help to reduce the impact of glass waste on the environment. From those three steps, the quite effective step in reducing the impact of glass waste is recycle stage. Recycle method is the process of turning a used material into a new material to prevent waste. By applying the recycling process, it is expected that glass waste can be utilized as reusable materials and can reduce waste at a certain limit, can save natural resources and reduce dependence on certain raw materials.

The research reported by Celik [7] on ceramics making using mineralogy method and chemical analysis method which is a sample of raw soil was carried out using X-ray diffraction (XRD) techniques to obtain clearer soil patterns. Clay minerals can be expanded and then heated to 550ºC for 3 hours to differentiate chlorite and kaolinite. The XRD pattern is obtained from Rigaku Rint-2200 diffractometer that operates at a tube voltage and current of 40kV and 30mA. Related to "Thermal analysis (DTA-TGA)", Netzsch STA 409 differential calorimeter tool under atmospheric air is used to analyze termal termo gravimetri termal (DTA-TGA). The temperature is raised from room temperature to 1200° C at a speed of 10° C per minute and maintained at this maximum temperature for 10 minutes. In "Dilatometry test", raw clay is heated to 1200° C at a heating rate of 50º C per minute to 1100º C, and 30ºC per minute for temperatures between 1100º C and 1200° C is maintained at a maximum temperature of 10 minutes. Samples that are dried overnight at 105º C were then heated in a Misura dilatometer of 3.32 (ODHT-HSM) in which 1600/80 changes are recorded every minute during heating. From the research, it is concluded that clay from Afyon Instambul is characterized using chemical, mineral and thermal analysis. Afyon clay has the qualities needed to make ceramic tiles and is an alternative ceramic raw material for developing Turkish ceramic tiles.

In [8], the authors reveal that alumina-oriented ceramics with fraction textures ranging from 9.6% to 93.6% are prepared by templated grain growth (TGG) of nanoscale matrices, the fraction texture can be controlled directly on ceramic density, and absorption of grain texture can be observed at ≥ 58.4%. The research used the X-ray diffraction (XRD) method, morphological features can be observed by scanning electron microscopy (SEM), to measure the density using the Archimedes method, the characteristics of mechanical properties are measured by three bends in universal testing, while hardness testing uses Vickers indenter (HVS-5). By this, it is concluded that the microstructure, mechanical properties, and the development period of alumina ceramic textures occur the compaction in the sample. Anisometric granules attachment occurs when orientation of ≥ 58.4%, cracked deflection between surfaces originating from oppressed grains because energy fractures is lower than the surface, in which further enhancements result in longer cracking deflection distances, texture characteristics of 93.6% alumina ceramics substantially increase Kc = 4.6 MPa.m1/2 and σf = 589 MPa, and therefore the luck occurs when the crack deflection.

In [9], the researchers report that the implementation of innovative materials for energy-saving focusing on the building sector is a very interesting topic throughout the world. In this case, aerogel is indicated as a promising material because it has increased in recent years in which aerogel has very low thermal conductivity around 0.01-0.02 W/(m.k). In his paper, he explained that the development of a dual glass panel system is with monolithic silica aerogel. It is mentioned that the results of the multi-disciplinary analysis include thermal and lighting tests. From the results, it is concluded that research carried out in different climates, the aerogel window can be used as a solution to reduce energy demand and even the demand in winter by reducing the increase in solar heat in buildings. It is important to state that when the internal burden on school buildings and office buildings is high, climate adjustment tends to be considered. And moreover, the use of aerogel glass on the window allows a lot of light during the day enters the room which affects on activity and prosperity. The window glass in this study is applied in real life, especially at the risk of condensation and heat changes.

In [10], the authors pinpointed that the characteristics of glaze sanitary physical properties are controlled and improved by changing the chemical composition of raw materials. Glaze industry from ten raw materials has different compositions, by spraying traditional ceramic substrates, then thermally treated and processed at 1250ºC. The glaze characteristics obtained are diffracted with X-rays to reveal the mineralogical composition and are confirmed by the FTIR spectrum and observed through images to see the electron distribution to study the microstructure and then, it is found that micro characteristics are obtained in white from the colorimeter data. From the results, it is concluded that the 10% difference in calculations and experiments is derived from the coefficient of thermal expansion and white color increase reaches 87.5%, the maximum zircon content is 14.5wt%, the lowest ZnO content is 2.5wt% and the flexural strength is 47.61 MPa.

The same research has also been reported by Martinez et. al [11] about the development of glaze material with glass and calcite-based material ($CaCO_3$). In this research, variations in holding time are carried out including 15 minutes, 30 minutes and 45 minutes with a combustion temperature of 1000ºC using a furnace machine and also the hardness test (Vickers) is carried out. Then, it is concluded that the micro photo showed that at 15 minutes holding time, there is cracking on the ceramic glaze layer, 30 minutes holding time has also cracking but there is a pinhole in the glaze layer, while at the 45 minutes holding time, the pinhole and cracking decreases. In the hardness test, the 15 minutes holding time is 452,761 VHN, the 30 minutes holding time is 479,914 VHN, the 45 minutes holding time is 391,695 VHN. This is done successfully and it obtained the highest results at a holding time of 30 minutes.

According to Demidenko [12], the characterization of calcium silicate ceramics using scanning electron microscopy (SEM) techniques sintered at temperatures of 850ºC to 1100ºC reveals that the microstructure of calcium silicate ceramics is going better at sintering temperatures of 1100ºC with a density of 2.32 g/cm^3, and porosity of 27.5%.

In [13], the authors examine the zircon corrosion behavior of glazed ceramic glass against strong acid solutions. It showed that the composition of the glass phase is determined by its glaze resistance. When ZrO_2 is involved in the glassy phase, very poor resistance is obtained through the formation of soluble Zr-containing compounds. Enhancing SiO_2 content is enhanced the resistance of the glass phase and dissolution of the observed grain or grain boundary phases. The coalescing of ZrO_2 into the glass phase must be avoided to increase the chemical resistance of zircon containing glaze. When ZrO_2 is in the grains, good chemical resistance can be obtained.

Additionally, In [14], the authors concluded that the mixture chosen from waste and low-cost mineral initial materials can be converted into glass ceramic components either by vitrification and subsequent crystallization or by direct sintering. The sintered-crystallization approach leads to strong glass ceramics, under the simple conditions (pressing fine powder and sintering at 950ºC for only 30 minutes). The porosity of the glass-ceramic body surface from direct sintering can be sealed by ceramic glaze which resulted from the same initial mixture of waste and minerals. The characteristics of the waste ceramic glaze can be adjusted by the addition of secondary components (glass panels from disassembled CRT and zircons). Because of their specific strength characteristics, coated ceramic glass can be applied in the building industry as lightweight tiles.

In [15], the authors reported that the milling method used is a technological scheme for the synthesis of ceramic materials using waste material from oil refineries containing heavy metals. This study experimented with the SEM test and revealed that pigments calcined in experiments using temperature of 700, 900, has a morphology that is very agglomerated, while, at a temperature of 1100, agglomerates that have formed undergo structural decomposition. For compositions containing almost all metal iron (92%) turned to magnesium chromium at temperatures of 1000. It is concluded that the proposed alternative technology scheme could provide economical production of ceramic pigments with exceptional compositional stability at relatively low temperatures from calcination.

In [16], the authors reported that glass powders are used as an alternative material to replace cement. In this study, glass powders are as an additive to cement because it depends on the chemical elements present in glass and also in cement, namely silica (SiO_2). In this study, it is compared using four mixed compositions, among others 0%, 10%, 20%, 30% with a total of seven specimens each with a 28-day treatment period based on SNI 03-0349-1989 standards. The results indicated that the use of glass powder on bricks with a mixture composition of 10%, 20%, 30% has met the requirements for the water absorption of solid concrete brick quality level I according to the provisions of SNI 03-0349-1989, bricks with a composition of 0% are included in quality level III according to the provisions standard of SNI 03-0349-1989 with an average compressive strength of 58.16 kg/cm^2, while bricks with a composition of 10%, 20%, 30% are included in quality level II with an average compressive strength of 74,41 kg/cm^2. So, it can be concluded that the most optimum mixture is bricks with a mixture composition of 20%.

In [17], the authors explained that the heat treatment process is an effort to increase the strength and hardness of steel by heating the steel to achieve austenite temperature that is followed by quench so that the martensite phase arises. In principle, the surface treatment of the material is almost the same but only done on the surface of the material in which the main goal is to get a hard surface of the component but the inside of the component remains resilient. From that condition, the method used is induction heating. Principally, it uses heating obtained from Eddy currents caused by flux magnetic originating from the coil of alternating electric current. Furthermore, it is tested using induction heaters for the heat treatment process on the surface of ST 37 steel specimens with the room temperature controlled by thermodigital, while, the specimens are quenched using a water cooling medium.

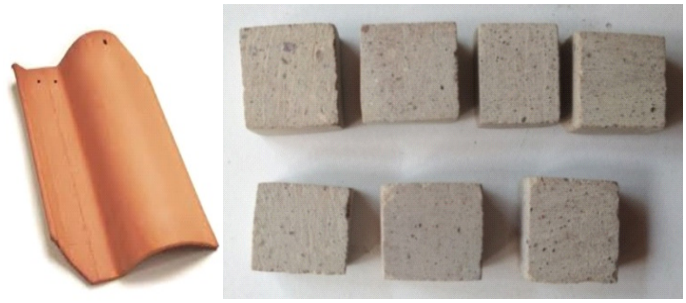
The manufacture of quality hardflek products from plastic waste and rice husks. The method of processing plastic waste and rice husk into hardflek products as a substitute for wood raw materials in furniture products has a high creative value in which quality and productivity is much better. Furthermore, the economic value of a furniture product is largely determined by the design or shape and appearance offered by the raw material. Plastic waste material combined with rice husk waste provides a very attractive aesthetic nature with interesting color because the coloring has been integrated in the material so that it cannot fade or be degraded by time. Processing methods of products made from polyethilene terepthalate (PET) plastic waste is reinforced with rice husk waste. The method applied can use the system hand lay-up (pouring). The test results showed that the addition of rice husk ash at a volume fraction of 93% plastic waste and 7% rice husk ash has a lower impact strength value than that of the initial specimen (without rice husk) from 0.0024 joules/mm^2 to 0.0015 joules/mm^2. The hardness testing of the Vickers method showed that the addition of rice husk ash in a volume fraction of 93% plastic waste and 7% of rice husk ash has increased from the initial specimen which is about 35% (10.6 HV to 13.5 HV), so there is a need for research development with SEM (scanning electron microscopy) method to know the composition and properties of the material contained in the product [18].

All things considered, it was necessary to conduct research that utilizes used bottle waste to be used as a more useful material in the form of glaze material. The purpose of this research was to develop glaze material by utilizing glass waste from used bottles combined with kaolin. To find out the characteristics and structure of glaze material, it was conducted the micro photo testing and hardness level analyzing of the glaze material using the Vickers hardness test.

2. Research Methodology

2.1. Materials

This study used bottle waste as the main material to coat the surface of the tile and also kaolin as an additional material in the coating process in which the combustion process used a furnace machine. Three types of time variations used were 30 minutes, 45 minutes and 60 minutes, besides, variations comparison of material composition between glass and kaolin were also used by the comparisons of 90:10, 80:20 and 70:30. This experiment began with the preparation of tile material since the tile material will be coated, followed by cutting the tile with a length of 2 cm, width 2 cm, and thickness 0.8 cm. This process used tile material that had not been burned. Figure 1 showed the tile material that had been cut and sanded using sandpaper ranging from 800, 1000, 1500, 2000 to 5000 grits.



(a) (b)

Figure 1: Clay tile specimens: (a) before being cut, (b) after being cut and mashed

Glass pounding machine was used to refine glass from used bottles into powder with a size of 200 meshing. To refine the used bottle glass used a glass pounding machine in an existing glass pulverizer at the Metallurgy Laboratory. Then it was meshed using size 200 mesh. Figure 2 shows the glass powder that had not been meshed and after meshing.

Kaolin (clay) is a clay mineral containing several layers of aluminum silicate. Kaolin is a kind of clay that is smooth, soft and white resulted from weathering granite rocks which is then used as material to make porcelain or to make a mixture of woven fabric (paper, rubber, and medicine, etc.), commonly called Chinese clay. Furnace machine is an equipment used to heat materials and change their shape (eg rolling, forging) or changing its properties (heat treatment). It is commonly referred to as an oven. Energy transfer in the furnace occurred during the heat energy by generation steps through element heater whose energy was supplied from electrical energy. In this study, a furnace machine was used to heat glass powder on a tile.



(a) (b) (c)

Figure 2: Recycle bottle glass (a), Glass powder before meshing (b) and after meshing 200 mesh (c).

The coating process is done manually with a thickness of 1mm. It is done by giving the liquid mixture of glass powder, kaolin, and water on the tile. This to mix glass and kaolin powder so that it can stick on the surface of the tile. The drying process is carried out so that the coating adheres firmly to the specimen. This process is done at room temperature and allowed to stand for a day. The combustion process uses ceramic tile specimens that have been coated with glass powder and kaolin. This process is carried out using a furnace engine involving necessary steps which are when the furnace engine is turned on, set the temperature on the screen with 1000°C, wait for the furnace machine to show the temperature of 1000°C on the screen in front. When the temperature already reaches 1000°C, set the time with a stopwatch for 30, 45, and 60 minutes. When the stopwatch shows the desired time, turn off the furnace machine and leave the specimen in the furnace machine until the temperature drops using room temperature. After the furnace machine reaches room temperature, then take a specimen from the furnace machine, then leave it for a day to proceed to testing.

2.2. Experiments

In the research, to investigate the influence of kaolin contents and holding time to the glaze hardness the micro hardness test was conducted. The experiment is carried out based on the ASTM E384 standard. The force was 0.5 kgf and 10 s of holding time.

Micro-photograph analysis testing was conducted to determine the microstructure of the results of the coating glass process on the tile. The observation is done by obtaining an image that shows the distribution of ceramic material with a zoom of 200 times, and also by arranging the lighting. From this picture, it finds out the properties and characteristics of ceramic materials.

3. Results and Discussion

3.1. Photo Micro

To facilitate reading comprehension in this test analysis, the glaze coating carried out using a variety of composition compared to glass powder and kaolin consisted of (A) 90% : 10%, (B) 80% : 20%, and (C) 70% : 30. There was also a variation of the time ratio performed including 30 minutes, 45 minutes, and 60 minutes, with the same temperature of 1000ºC. This was to find out which one is better for use after the rain process.

Figure 3 shows the results of micro-photographs on the surface of the glaze material. Figure 3 is the results of micro-photographs with the variations in the composition of the glaze material with a comparison between glass and kaolin.

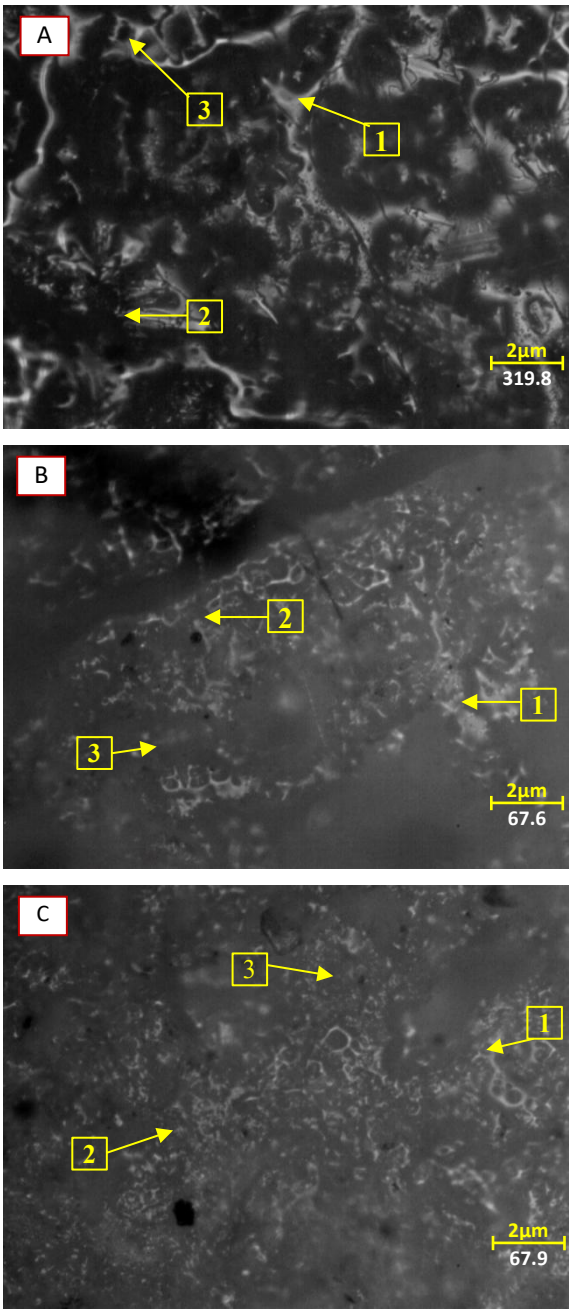


Figure 3: The results of micro-photographs on the glaze material holding time 30 minutes

From the observation results of micro-photographs tests conducted in Figure 3 (A), number 1 is shown in the form of longitudinal lines like fibers and marked in white. It is the melted glass material and has formed on the glaze layer on the surface of the tile or commonly called amorphous solids. While the black one shown in number 2 is porosity. It can possibly occur when kaolin as a binding material glaze is burning to ash and trapped on the surface of the glaze layer. Then, number 3 shows that the existence of pinholes that are possible to occur due to the release of oxygen during combustion and inside the material contains flammable materials at low temperatures. The oxygen arises in the combustion process because the room in the furnace is not in a vacuum condition, so the materials are combusted first and form holes on the surface of the layer. The same process is also shown in the figure B and figure C in which number 1 shown as the shape of a long line that looks like a fiber and marked white is a glass material that has melted but is somewhat less than what is shown in figure A. It is possible occurs because it affects the material composition, whereas porosity that occurs in figure B and C looks more possible occurs because it has more kaolin. The pinholes shown in figure B and C shows that those are in glaze layer and have a smaller size with a greater amount than in figure A. Figure 3 analysis shows that the use of temperatures of 1000ºC during the combustion process successfully melt the glass.

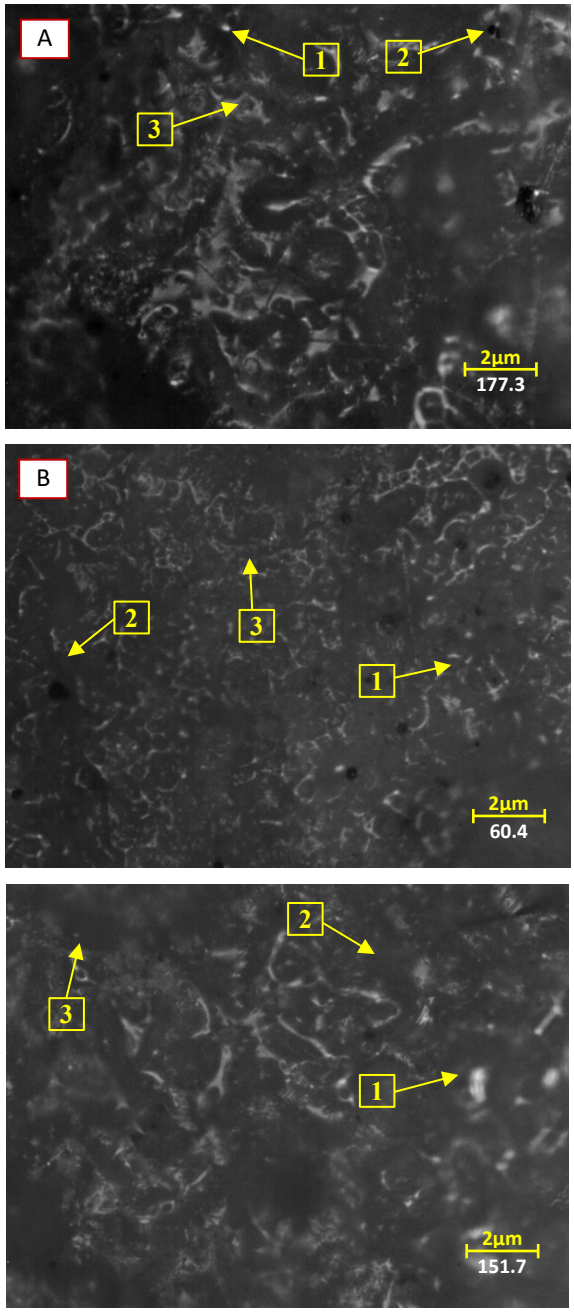


Figure 4: The results of micro-photographs on the glaze material holding time 45 minutes.

Figure 4 shows the results of micro-photographs on the surface of the glaze material. Figure 4 is the result of micro photographs with composition variation of the glaze material with a comparison between glass and kaolin. The longitudinal lines like fibers that are

marked in white are melted glass materials or commonly called amorphous solids. It forms a glaze layer on the tile surface, but it can be seen that the melting rate of glass material is not evenly distributed well. Equally, it happens in figure B and C. It can be seen that the melting rate of glass material that occurs in figure B shows that the melting rate of glass material looks more evenly distributed if compared to the figure A and C. This is possible due to differences in the composition of glass and kaolin. Figure A number 2 shows porosity in which it occurs in smaller and fewer sizes. Porosity also occurs in figure B and C, but in figure C the porosity looks less if compared to figure A and figure B. Pinholes that occur in the observation can be seen in number 3, but in part B, it looks smaller pinhole. This possibly happens because it depends on the length of time to hold the combustion and composition comparison.

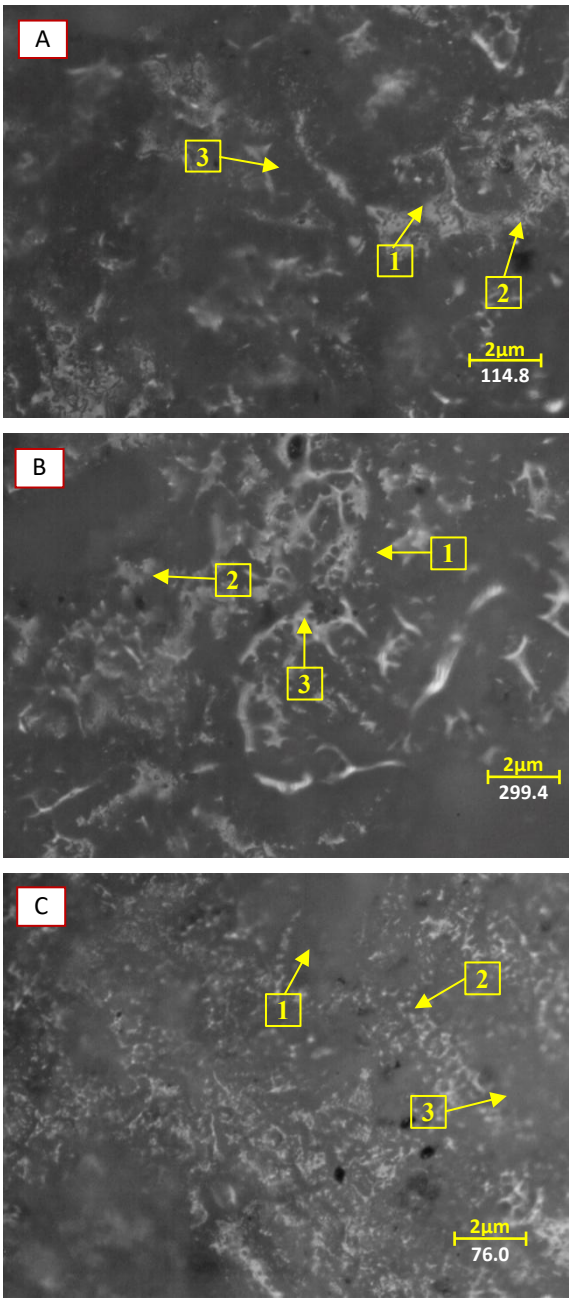


Figure 5: The result of micro-photographs on the glaze material holding time 60 minutes

Figure 5 shows the results of micro-photographs on the surface of the glaze material. Figure 5 is the result of micro-photographs with variations in the composition of the glaze material with a comparison between glass and kaolin. The observation results of micro-photographs tests conducted in figure A, number 1 is shown in the long lines like fibers and marked in white. It is the melted glass material and has formed on the glaze layer on the surface of the tile or commonly called amorphous solids. But it can be seen that the melting rate of glass material is not evenly distributed properly. Identically, it happens in figure B and C. But it can be seen that the melting rate of glass material that occurs in figure C shows that the melting rate of glass material looks more evenly if compared to figure A and B. This is possible because of differences in the glass and kaolin composition. The observations can be strengthened by other studies [19][2]. Number 2 shown in figure A, B and C shows the porosity, but the porosity seemed less is found in figure C compared to figure A and B. This possibly happens because it affects the composition variation [2]. Pinholes are shown in number 3 of figure A, B and C. In figure A, there is no pinhole compared to figure B and C. This may occur due to differences in composition in figure A and the length of combustion duration.

From the analysis results of micro-photographs on glaze material with a holding time of 30 minutes, 45 minutes and 60 minutes, as well as the variations of material composition comparison between kaolin and glass powder with a ratio of 90:10, 80:20, 70:30, figure 5 part C that has a variation of 70:30 and a holding time of 60 minutes shows the results of a dense glaze layer compared to other variations. It is because the glaze materials have different compositions and holding times.

3.2. Micro Hardness

Figure 6 is the result of the photos of glass material that has been coated and has been through the heating process with the furnace. Then, it will conduct the process of specimen testing of finished material using a hardness test tool.

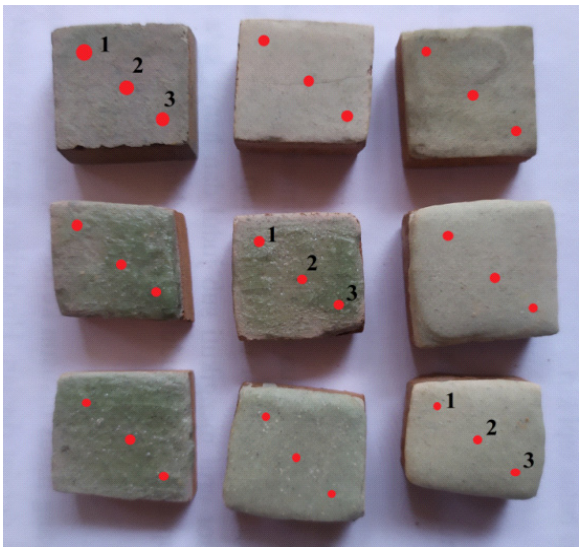


Figure 6: Location of the indertor point on micro Vickers hardness testing

Specimens used in this test were tile materials that had been coated with glass powder and kaolin with variation composition after being heated at 1000º C. The combustion aims at increasing

the hardness on coated material and the surface a thin layer was formed. After completing the combustion process, a hardness test (Vickers) is conducted to find out the hardness value on the surface of each material. The rectangular diamond pyramid indenter is used in the hardness test with 0.5 kgf and a time of 10 seconds is set as the loading test. The results of the hardness test that have been carried out can be seen in figure 7.

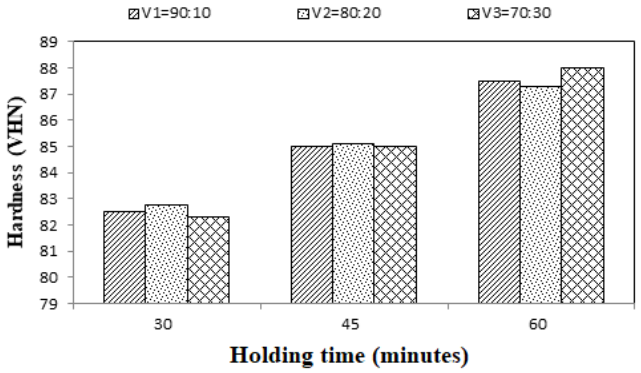


Figure 7: Comparison of glaze hardness values with composition variations

The overall results of the hardness test figure 7 analysis is a variation of the glaze composition between glass and kaolin, and different variations of the holding time. It has been proven by test results in figure 7. It can be stated that the process of a hard layer forming on the surface of tile material that is on the highest Vickers hardness testing occurs at a holding time of 60 minutes and composition variation of 70: 30% with an average hardness value of 88.1 VHN. While the lowest hardness is at a time variation of 30 minutes with an average hardness value of 81.9. The difference in the hardness value in each variation is possible because the characteristics of microstructure formed are also different. This is evidenced in section 4.1 which discusses the results of micro-photographs testing in which the less of the diminishing product defects in the form of pinhole and porosity, the higher the hardness.

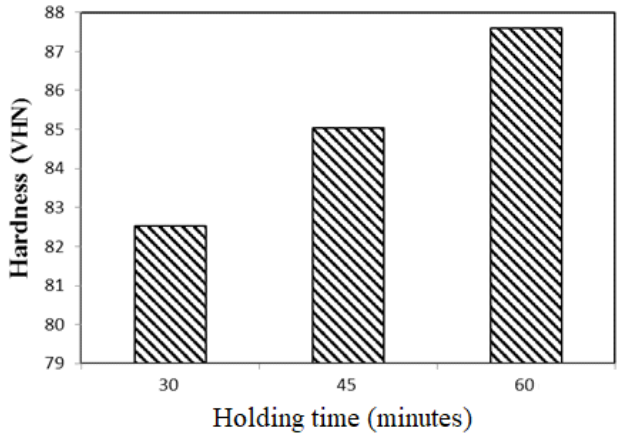


Figure 8: Comparison between the average hardness value (Vickers) with time variations.

4. Conclusion

From the results of the analysis and discussion regarding the coating of glaze material using glass and kaolin based materials, it can be concluded that the time variation is very influential on the results of micro-photographs analysis and hardness test. So, the longer the holding time given, the more the melting rate of the glass evenly distributed and the hardness test has also increased. The results of micro-photographs analysis on glaze material samples using glass and kaolin base materials are combusted at a temperature of 1000ºC with different variations of holding time of 30 minutes, 45 minutes, and 60 minutes. It can be concluded that the holding time of 30 minutes indicates the glass melt at the holding time of 30 minutes maximum melting at this temperature, and there are still pinholes and porosity in the analysis. At the 45 minutes holding time, it shows that the glass melt has not been evenly distributed since there are some pinholes and porosity but it looks less and smaller than the 30 minutes holding time. At the 60 minutes holding time, the glass melt is still visible but looks smaller and less, porosity and pinhole are still visible but look smaller and less than the holding time of 30 minutes and 45 minutes. The results of hardness test (Vickers) analysis revealed that the hardness value increases in each composition variation carried out. At a holding time of 30 minutes, 45 minutes and 60 minutes, it shows the increase average hardness value of 82.53 VHN, 85.03 VHN, 87.6 VHN.

Conflict of Interest

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Acknowledgment

Authors would like to thanks to Universitas Muhammadiyah Surakarta, and Master Program of Mechanical Engineering for supporting the research.

References

[1] M. J. Davis and E. D. Zanotto, "Glass-ceramics and realization of the unobtainable: Property combinations that push the envelope," *MRS Bull.*, vol. 42, no. 3, pp. 195–199, 2017. https://doi.org/10.1557/mrs.2017.27

[2] J. Deubener, M. Allix, M.J Davis, A. Duran, T.Hoche, T. Komatsu, S. Kruger, I. Mitra, R. Muller, S. Nakane, M.J. Pascual, J.W.P. Schmelzer, E.D. Zanotto, S. Zhou., "Updated definition of glass ceramics," *J. Non. Cryst. Solids*, vol. 501, no. January, pp. 3–10, 2018. https://doi.org/10.1016/j.jnoncrysol.2018.01.033

[3] M. Rampf, M. Dittmer, C. Ritzberger, and W. Höland, "Controlled parallel crystallization of lithium disilicate and diopside using a combination of internal and surface nucleation," *Front. Mater.*, vol. 3, no. October, pp. 1–9, 2016. https://doi.org/10.3389/fmats.2016.00047

[4] A. Halliyal, A. Safari, A. S. Bhalla, R. E. Newnham, And L. E. Cross, "Grain-Oriented Glass-Ceramics for Piezoelectric Devices," *J. Am. Ceram. Soc.*, vol. 67, no. 5, pp. 331–335, 1984. https://doi.org/10.1111/j.1151-2916.1984.tb19532.x

[5] V. Luntz-Leybman, R. K. Freund, and A. C. Collins, "5A-Pregnan-3A-Ol-20-One Blocks Nicotine-Induced Seizures and Enhances Paired-Pulse Inhibition," *Eur. J. Pharmacol.*, vol. 185, no. 2–3, pp. 239–242, 1990. https://doi.org/10.1016/0014-2999(90)90648-P

[6] Yasin Erdoğan, "Physicochemical Properties of Handere Clays and Their Use as a Building Material", Journal of Chemistry, Volume 2015. http://dx.doi.org/10.1155/2015/374245

[7] H. Celik, "Technological characterization and industrial application of two Turkish clays for the ceramic industry," *Appl. Clay Sci.*, vol. 50, no. 2, pp. 245–254, 2010. https://doi.org/10.1016/j.clay.2010.08.005

[8] M. Zhang, Y. Chang, R. Bermejo, G. Jiang. Y. Sun, J. Wu, B. Yang, W. Cao., "Improved fracture behavior and mechanical properties of alumina textured ceramics," *Mater. Lett.*, vol. 221, pp. 252–255, 2018. https://doi.org/10.1016/j.matlet.2018.03.123

[9] U. Berardi, "The development of a monolithic aerogel glazed window for an energy retrofitting project," *Appl. Energy*, vol. 154, pp. 603–615, 2015. https://doi.org/10.1016/j.apenergy.2015.05.059

[10] K. Boudeghdegh, V. Diella, A. Bernasconi, A. Roula, and Y. Amirouche, "Composition effects on the whiteness and physical-mechanical properties of traditional sanitary-ware glaze," *J. Eur. Ceram. Soc.*, vol. 35, no. 13, pp. 3735–3741, 2015. https://doi.org/10.1016/j.jeurceramsoc.2015.05.003

[11] Juan Daniel Martínez, Santiago Betancourt-Parra, Ivonne Carvajal-arín, Mariluz Betancur-Vélez, "Ceramic light-weight aggregates production from petrochemical wastes and carbonates (NaHCO3 and CaCO3) as expansion agents," . Construction and Building Materials, Volume 180, 2018, Pages 124-133, https://doi.org/10.1016/j.conbuildmat.2018.05.281.

[12] N. I. Demidenko and G. B. Tel'nova, "Microstructure and properties of a material based on natural wollastonite," *Glas. Ceram. (English Transl. Steklo i Keramika)*, vol. 61, no. 5–6, pp. 183–186, 2004. DOI: 10.1023/B:GLAC.0000043088.65135.11

[13] G. Topateş, B. Tarhan, and M. Tarhan, "Chemical durability of zircon containing glass-ceramic glazes," *Ceram. Int.*, vol. 43, no. 15, pp. 12333–12337, 2017. https://doi.org/10.1016/j.ceramint.2017.06.097

[14] M. A. Binhussain, M. Marangoni, E. Bernardo, and P. Colombo, "Sintered and glazed glass-ceramics from natural and waste raw materials," *Ceram. Int.*, vol. 40, no. 2, pp. 3543–3551, 2014. https://doi.org/10.1016/j.ceramint.2013.09.074

[15] M. Doynov, T. Dimitrov, and S. Kozhukharov, "Alternative technological approach for synthesis of ceramic pigments by waste materials recycling," *Bol. la Soc. Esp. Ceram. y Vidr.*, vol. 55, no. 2, pp. 63–70, 2016. https://doi.org/10.1016/j.bsecv.2016.01.002

[16] Fernanda Andreola, Luisa Barbieri, Isabella Lancellotti, Cristina Leonelli, Tiziano Manfredini, "Recycling of industrial wastes in ceramic manufacturing: State of art and glass case studies", Ceramics International, Volume 42, Issue 12, 2016, Pages 13333-13338. https://doi.org/10.1016/j.ceramint.2016.05.205.

[17] Anggono, A. D., Muttaqiem, Z., Darmawan, A. S., Besar Riyadi, T. W., Yulianto, A., Sugito, B., Abdullah, E. M. (2019). Mechanical Properties of Concrete Block Reinforced with Recycle HDPE and Coal Bottom Ash. Materials Science Forum, 961, 51–56. https://doi.org/10.4028/www.scientific.net/msf.961.51

[18]. A. D. Anggono, B. Sugito, A. Hariyanto and Suranto, "Investigation on mechanical properties of friction stir welding of 2 mm thick aluminium alloy sheet," 7th Brunei International Conference on Engineering and Technology 2018 (BICET 2018), Bandar Seri Begawan, Brunei, 2018, pp. 1-4. doi: 10.1049/cp.2018.1519

[19] X. Lu, J. Yang, X. an Ning, K. Shih, and F. Wang, "Crystallization pathways in glass-ceramics by sintering cathode ray tube (CRT) glass with kaolin-based precursors," *J. Eur. Ceram. Soc.*, vol. 38, no. 15, pp. 5184–5191, 2018. https://doi.org/10.1016/j.jeurceramsoc.2018.06.047

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 317-324 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Stress Response Index for Traumatic Childhood Experience Based on the Fusion of Hypothalamus Pituitary Adrenocorticol and Autonomic Nervous System Biomarkers

Noor Aimie Salleh*,[1], Malarvili Balakrishnan [1], Anna C. Whitttaker [2]

[1]*School of Biomedical Engineering and Health Science, Faculty of Engineering, Universiti Teknologi Malaysia, Johor Bahru Malaysia*

[2]*Faculty of Health Sciences and Sport, University of Stirling, Stirling, United Kingdom*

ARTICLEINFO

Article history:
Received: 30 September, 2019
Accepted: 19 January, 2020
Online: 07 February, 2020

Keywords:
Heart Rate Variability
Salivary Cortisol
Stress
Biomarkers
Fusion

ABSTRACT

Stress occurring in the early days of an individual was often assumed to cause several health consequences. A number of reports indicated that having to deal with unfavourable events or distress situation at a young age could tweak stress responses leading to a broad spectrum of poor mental and physical health condition. Therefore, changes identified within stress response were recommended to be taken as a measure in regulating and managing such health situation. This study combines the biomarker that represents both autonomic nervous system (ANS) and hypothalamic-pituitary-adrenocorticol (HPA) as a single measure to classify the stress response based on traumatic childhood experience and propose a stress response index as a future health indicator. Electrocardiograph (ECG), blood pressure, pulse rate and salivary cortisol (SCort) were collected from 12 participants who had traumatic childhood experience while the remaining 11 acted as the control group. The recording session was done during a Paced Auditory Serial Addition Test (PASAT). HRV was then computed from the ECG and the HRV features were extracted. Next, the best HRV features were selected using Genetic Algorithm (GA). Biomarkers such as BP, PR and SCort were then integrated with 12 HRV features picked from GA. The integrations were conducted using two fusion methods which are Euclidean distance and serial fusion. The differences in reaction of the fused features were then identified. Based on the result, the Euclidean distance (ed) which is the fused feature by the parallel fusion, displayed the most efficient reaction with accuracy, sensitivity, and specificity at 80.0%, 83.3% and 78.3%, respectively. Support Vector Machine (SVM) was utilized to attain such result. The fused feature performance was then fed into SVM which produced indexes on stress responses. The result retrieved from these indexes acts as a measure in handling future health deliverability and perhaps could eventually enhance the health care platform for midlife individuals.

1. Introduction

This paper is an extension of the work [1] originally presented in the IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), 2018. Stress is identified as a biological and mental reaction to any threat or external pressure. During normal conditions, stress encourages a person to concentrate, stay active and attentive. Nevertheless, in critical events, stress assists in powering up an individual and this may lead to the ability in saving someone's life. However, stress becomes damaging to the physical and psychological health of an individual when the body fails to cope and solve problems [2]. Stress can initially be a positive contributing factor in an individual's life but due to immoderate demands that complies with the advanced and modern world eventually becomes a major threat. Many acute diseases have been listed as the effect of stress. This includes cardiovascular diseases, diabetes, asthma, anxiety disorders and depression [3].

Stress is classified into three types which are known as acute, episodic and chronic [3,4,5]. Acute stress is known as stress occurring at a short period of time. Factors contributing to this kind of stress are usually work-related, athletic performances and examination. Episodic stress occurs due to repetitive acute stress, like that of a daily work stress. Chronic stress is related to a high intensity stress factors that develops over a long period of time. This

*Corresponding Author: Noor Aimie Salleh Email: aimie@biomedical.utm.my

can be damaging to the physical and psychological well-being of an individual [6]. Factors that can lead to development of chronic stress are marital affairs, work stress, financial condition [7], and traumatic life experiences [8, 9] such as an unpleasant childhood incident.

Several researchers have been caught up studying stress factor developed from traumatic childhood experiences. Recent studies in illness management and prevention indicates that being stressed at an early age may lead to a series of health issues [10, 11, 12, 13], and alterations in physiological behaviour [14, 15, 16, 17]. This fact is significant in developing methods that could improve healthcare management for individuals approaching adulthood, since there is an indicator of the current health status of the particular child. Therefore, this study focuses on traumatic childhood experience which is classified as chronic stress. In sequence to investigate the stress response of individual with traumatic childhood experience, mental stress test which is an acute stress is used.

Early life stress has shown a number of health outcomes to be predictive [18, 11]. Studies conducted a few years ago showed that individuals with unfavourable and stressful situation tend to fall in a broad spectrum of mental and physical health problems which eventually lead to alcohol abuse, obesity, depression, smoking, substance abuse, intrusive sexual acts [19], as well as acute illnesses like cardiovascular disease [20, 21]. The output of these studies has brought latest discovery that indicate children who are vulnerable to potentially traumatic events may experience cardiovascular stress reactivity due to acute physiological processes. There are a number of cases that displayed changes in cortisol as well as severe mental stress that increases heart rate within individuals who had at least one or more traumatic incident [14, 15]. Nevertheless, studies also showed changes in cardiovascular reactivity to stress task amongst children who were exposed to aggressive adolescent behaviours [17]. These evidences indicate that stress reactivity is detected via incidents taking place at an individual's tender age, and that the immensity of stress reactivity is in relation to stress exposure and may influence health of an individual in years to come [22, 23]. Since there are compelling evidences on this matter, detecting changes in stress response due to traumatic events of a child is significant to be utilized as an indicator for better healthcare.

It is important to pick the most suitable biomarker to identify and determine changes taking place in stress response. Heart rate variability (HRV) is one of the biomarkers proven to be the most powerful and depicts autonomic nervous system [24-27]. However, there is not much research on the application of HRV in measuring stress responses among children with traumatic incidents. In fact, only conventional frequency domain method was applied [28, 29]. Hence, both conventional and improved approaches extracted from HRV analysis should be applied in measuring stress responses of a child who had been through traumatic incidents.

Integrating two of the major body systems known as autonomic nervous system (ANS) and hypothalamic-pituitary-adrenocorticol (HPA) is strongly recommended as it plays an important role in regulating stress response [30,31]. Although numerous studies indicate that in measuring stress responses, multiple biomarkers were used. In conjunction to that, the results for each biomarker were analysed singly [32, 33].

Therefore, this study proposes a latest fusion method for stress response classification on traumatic childhood occurrence. Since children's traumatic incident causes inconsistency in stress response, the health status during adulthood can be determined via a proposed stress response index which could indicate the future health condition and lead to preventative measurements as well as further diagnosis that may refine the healthcare management. The objective of this study is to integrate both HPA and ANS so as to classify stress responses among children who faced traumatic experiences along with a stress response index that would be able to indicate health statuses in the later years of these individuals.

2. Method

2.1. Participants

609 students were screened with lifetime adversity inventory based on C-DIS-IV items and Childhood Traumatic Questionnaire (CTQ) [34]. This group of participants is made up of 511 females and 98 males (age, M = 19.2, SD = 2.99 years). Participants who scored within 2 to 5 and 2 or higher on the lifetime adversity measure and CTQ, respectively, were identified as individuals with traumatic childhood incidents whereas those who scored 0 on the lifetime adversity measure and CTQ were made the control group [14, 34]. A total of 23 participants, 12 participants who had traumatic childhood experience and 11 participants who were from the healthy control group. These participants have no history of cardiovascular illnesses, incurable diseases or other acute infection, and a current endocrine or immune disorder. Participants' consent was obtained and the study was approved by the University of Birmingham Ethnics Committee.

2.2. Mental Stress Test

The psychological stress test in this research was conducted with the Paced Auditory Serial Addition Test (PASAT) and it usually takes about 10 minutes to complete. Participants were given a series of single-digit numbers and were requested to add the digit to the numbers they previously heard. They were to say their answers out loud.

2.3. Procedure

A standard three-lead ECG placement and blood pressure (BP) cuff, were attached to the participants in sitting position. The test began with a 10 minutes adaptation period followed by another 10 minutes of resting baseline phase. Three saliva samples were collected from every subject with the first sample obtained at the last minute of the baseline phase. 10-minute PASAT was then completed and saliva sample was obtained once again. The recovery phase took place for the next ten minutes followed by last saliva sample collection. Participants' BP and heart rate (HR) were recorded every two minutes, four times for each phase, whereas ECG readings were recorded throughout the test.

2.4. Saliva Sampling and Cortisol Assays

In order to obtain samples of saliva, participants had salivette dental swab positioned in their mouth which was gently chewed for one minute and placed in the salivette tube. Three stimulated saliva samples were obtained in this study whereby first sample via a 10-minute baseline phase. As for the second and third sample, it was obtained right after stress test and initially 10 minutes to task, respectively. The salivettes were then centrifuged for five minutes at room temperature and 3500 rpm before being aliquoted. It was then made to freeze at -20 degrees until the time to be assayed.

The assayed cortisol was then duplicated by an enzyme-linked immunosorbent assay (ELISA) via a commercial kit (DRG Diagnostics).

2.5. Pre-Processing: From ECG to HRV

HRV were being quantified via MATLAB software based on obtained and pre-processed ECG signals. A notch filter was used in eliminating the 50 Hz power line interference. Pan and Tompkin's algorithm were applied to identify the QRS waves database [35].

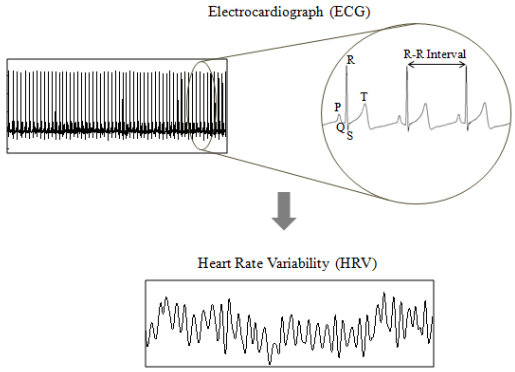


Figure 1: Derivation of Heart Rate variability for Electrocradiograph

2.6. Feature Extraction

In this research, HRV feature extraction is carried out through time, frequency, non-linear time-frequency and Wavelet analysis.

2.6.1. Time-Domain Analysis

The time-domain analysis was measured with Standard Deviation of the Normal-to-Normal intervals (SDNN), Standard Deviation of the Average of Normal-to-Normal intervals (SDANN), Root Mean Square Successive Difference (RMSSD) and HRV triangular index (HTI) [26] and computed via recommendation by Task Force (1996).

2.6.2. Frequency Domain Analysis

For the frequency domain analysis Autoregressive (AR) spectral analysis was applied [36,37]. AR model is interpreted as follow:

$$x[n] = \sum_{i=1}^{N} a_i x[n-1] + \varepsilon[n] \quad (1)$$

where, *x[n]* is the current value of the HRV time series, *a1, . . . , aN* are the predictor coefficients, *N* is the model order. With this method, the power spectral density (PSD) can be presented at a precise estimation. This would allow the feature extraction process to be carried out at ease as shown in Figure 2 [38].

Next, nine features were obtained via PSD extraction. These nine features are VLF (0.003-0.04 Hz), LF (0.040-0.15 Hz), HF (0.15-0.4 Hz), the standard unit (n.u.) of LF and HF (LFnu and HFnu), LF/HF ratio, Total Power (TP) [26, 38,39], log LF (lnLF) and log HF (lnHF)[40].

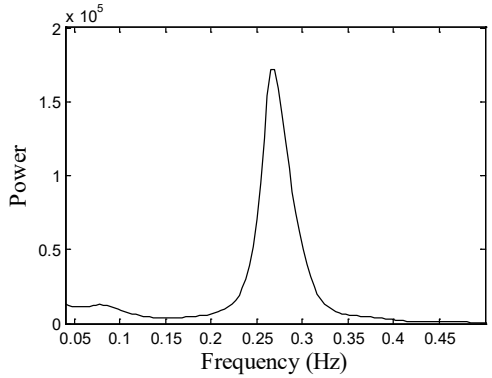


Figure 2: Power Spectral Density (PSD) using AR

2.6.3. Time-Frequency Domain Analysis

Since HRV contains slow varying signal, Modified B-Distribution (MBD) tend to be one of the most efficient Time Frequency Domains (TFDs) in exhibiting a high time-frequency resolution [41]. MBD's fundamental is [41]:

$$g(\nu,\tau) = \Gamma(\beta + j\pi\nu)^2 / \Gamma^2(\beta) \quad (2)$$

where, gamma function is identified as $\Gamma$and β is a positive real number between 0 and 1 that monitors the commutation between components' resolution and cross-term suppression [41].

Figure 3 shows an example of TF plotted via application of MBD to obtain HRV signals that were recorded during resting baseline. TFD measured via TFD-based Shannon and Renyi entropy led to an overall computation of TFD and as well as entropy within LF and HF domain [42,43].

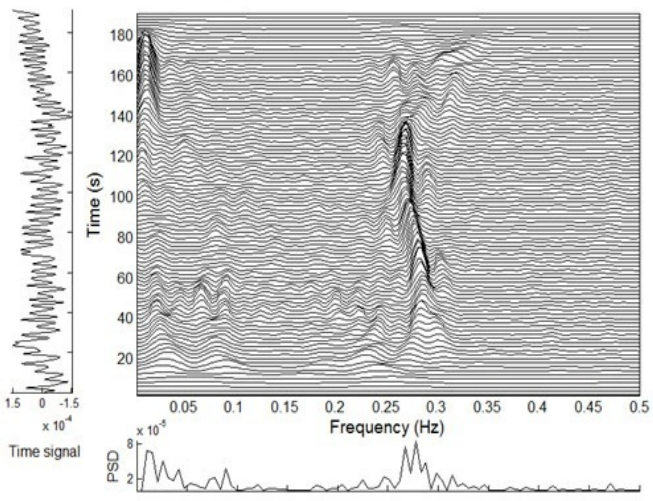


Figure 3: Time and frequency details via HRV's time-frequency distribution

2.6.4. Wavelet Transform Analysis

Wavelet analysis bring forth the localizations of time and frequency resulting in wavelet coefficients which can be utilized as features in classifiers [44,45]. HRV signal can be disintegrated using wavelet family $\psi a,b$, a basis function that involves dilatation and translations of a distinctive and valid mother wavelet $\psi(t)$. This process is defined as,

$$\psi_{a,b} = \frac{1}{\sqrt{a}}\psi\frac{t-b}{a} \quad (3)$$

where, scale and location is indicated by *a* and *b*, respectively [44, 45]. Basis function of DWT is represented at scale 2-m whereas, n indicates the time instant specified as follow,

$$\psi_{m,n} = \int_{-\infty}^{\infty} 2^{-\frac{m}{2}}\psi(2^{-m}t - n) \quad (4)$$

Hence, DWT's signal x(t) is specified as follow:

$$T_{m,n} = \int_{-\infty}^{\infty} x(t)\psi_{m,n}(t)dt \quad (5)$$

A discrete wavelet was transformed from eight-order Daubechies mother wavelet (db8) to five levels before being implemented to HRV. This transformation causes signal disintegration of five-level wavelet. Following this, using the suggested HRV frequency bands, Detail (D) and Approximate (A) coefficients were reassembled [26]. The A5 coefficient allowed rebuild of VLF range at < 0·04 Hz; D5 and D4 coefficients for LF range at 0·04–0·15 Hz; D23, D2, and D1 coefficients for HF range [46]. Hence, a total of 83 features were obtained from the HRV signal.

2.7. Feature Selection

The feature selection process is carried out via Genetic Algorithm (GA) MATLAB Toolbox. Iteration process occurred whereby chromosome is being engineered for production of new population via application of genetic functions like crossovers and mutations. Theoretically, healthy species will make through the evolution process whereas the weak ones disintegrate.

An initial population of chromosomes was randomly selected in order for GA's process to begin. Encoded bit strings usually depicts chromosome that is made up of sequence of features. These features each is classified as genes. In each chromosome, based on positional index, a gene value of '1' determines that the feature selected is positioned at '1' but the feature is not picked if the gene value is found to be at '0'. When iteration process takes place, combinatorial set of genes or features in the current population, identified as individual was observed via fitness function.

For the current work, the fitness function (*FitFunc*) was computed using kNN-based classification error with $k = 3$ [47,48,49] which is defined as:

$$FitFunc = \frac{\alpha}{N_f} + \exp\left(-\frac{1}{N_f}\right) \quad (6)$$

where, kNN-based classification error is indicated by α and Nf is the number of elements in selected features. In record, one study proposed the application of SVM classification error in extracting fitness function [50]. Unfortunately, despite being a substantial machine learning tool, SVM needs a durable computational load. Hence, in this research, this method was not applicable due to the constraint and fact that GA consumes computational space.

The initial populations of chromosomes were then ranked upon creation. The fittest individuals known as elitism solutions were selected from these ranks for survival in the next generation. Individuals that were not selected formed crossover and mutation solutions by genetically going through functional crossover and mutation. New generations were then formed by individuals selected through the elite, crossover and mutations.

GA operator's optimization was carried out to detect the prime value of initial population, crossover and mutation. Table 1 indicates the tested values extracted via references from studies carried out previously [47, 49, 51]. The outcomes of chosen features integrating with tested parameters were recorded and chromosomes with the optimum outcome were utilized in the next stage.

Table 1: Tested parameter figure utilized selection of HRV feature in GA.

Parameter	Tested Value
Initial Population	20
	30
	40
	50
	60
Crossover	0.1
	0.01
	0.001
Mutation	0.8
	0.9

Selection tool was required in choosing the best individual. Hence, in this research, two individuals or chromosomes were picked as defeater of the tournament by applying tournament selection of size 2. This selection was applied because of its clarity, rapidity and competency [47, 48]. The process is carried out in repetition until formation of new population is achieved.

Individuals like elite, crossover, and mutation kids were added to create the new population.

New Generation= Number (Elite kids) + Number (Crossover kids)…. + Number (Mutation kids) (7)

New Generation Score = Fitness (New Generation) (8)

This process continued until GA achieved the condition that puts it to stop. The stopping conditions utilized are Maximum Number of Generations (MaxNG) and Stall Generation Limit (SGLim). Both were set at 100 and 10, respectively [49]. If the mean alteration in the fitness gain between the chromosomes over SGLim generations was below than or equivalent to 0.000001 which is the value of Tolerances Function (TolFun), then the GA will be dismissed. The process ends if TolFun also known at termination tolerance, has a function gain whereby if $|f(xi) - f(xi+1)| <$ TolFun. Since GA observes the variance in values of fitness within all generations. The mean of these variances for 10 generations were recorded. If the recorded values were less than or equal to 0.000001, GA will be dismissed. This indicates that

resemblance in fitness value also known as genetic homogeneity among chromosomes of the generation have the most excellent chromosome and therefore coincides with GA.

2.8. Feature Fusion

In obtaining reactivity demarcation, the features were measure based on normalization equation as follow [52, 53]:

$$Y_n = \frac{Y_{stress} - \hat{Y}_{rest}}{\hat{s}_{rest}} \quad (9)$$

where, sample feature vector throughout stress texting is Y, the average of sample feature during resting baseline is indicated as $\hat{Y}$, and standard deviation of feature vector throughout resting baseline is indicated as $\hat{s}$.

Linking feature vectors of the biomarkers obtained from HRV feature extraction (h), and BP, PR and SCort (b, p and c, respectively) led to construction of serial fusion, whereas parallel fusion was designed by integrating feature vectors and complex vector. To introduce combined features with complex vector, the following equation, c = a + *i*b, where a and b are two varied feature vectors of the same sample S and *i* is an imaginary unit, was applied.

The purpose of fusion was to integrate the essentials within ANS and HPA leading to a finer presentation of the elements. Hence, the differences are obtained via fusion of biomarkers. ANS biomarkers are HRV, BP and PR, whereas HPA biomarker is cortisol.

2.9. Classification

In order to speculate an input's output group with existing data, a classifier is required. In this research, SVM classifier was used to classify stress response between two types of group which are of individuals with traumatic childhood incidents and a control group. Other classifiers like Naive Bayes [54, 55] and k-Nearest Neighbour (kNN) [56, 57] were used in this study to distinguish and prove that SVM is the most effective classifier. The evaluations of classifier were carried out by applying 10-fold cross validation [46].

2.10. Stress Response Index

The stress response index is depicted via classifiers that indicate greater deliverability. The future health is then predicted via irregularity patterns in stress response that were detected within the index.

3. Result and Discussion

3.1. HRV Feature Extraction

From the HRV feature extraction, a total of 83 features that has been extracted. Previous researches indicated that most features selected in obtaining HRV classifications were of conventional methods which are based on time and frequency analysis [58-60]. Nevertheless, there were studies that indicate the application of feature analysis that are more advanced such as that of wavelet, time frequency, and non-linear [46, 61, 62]. However, in this study, obtaining HRV features were carried out with the fusion of both conventional and newest advance analysis.

3.2. HRV Feature Selection

From genetic algorithm, 12 features were selected which are RMSSD from time domain analysis, Normalized unit of high frequency (Hfnu) from AR PSD, Mean LF, SDNN D1, SDNN D4, Apen (Approximate entropy) D4, SampEn (sample entropy) D3, Kurtosis D4, SkewD5, from wavelet and ShanEn LF, ShanEn (Shannon entropy) HF, ShanEn LFHFr from TFD.

3.3. Feature Fusion

The extracted HRV features via GA feature selection were then made to integrate with biomarker obtained in this study to produce stress response index. The initial biomarker fusion was carried out with the application of Euclidean distance (ed) classified as parallel fusion followed by serial fusion, a regular and simple fusion method, carried out for resemblance identification. Theoretically, the fusion of biomarkers that depicts ANS and HPA was applied in this research. This is because both ANS and HPA are biological systems with their own pathological characteristics and are strongly connected. Hence, a booming stress response index was likely to be extracted from the fusion of these systems [30, 31, 63]. Biomarkers depicting ANS were made to integrate with biomarker that depicts HPA, known as, Scort. With this fusion, HRV-Scort, BP-Scort and PR-Scort, were produced.

Table 2 shows the result of integrated biomarkers via the application of SVM in the context of accuracy, sensitivity and specificity. Outstanding performance can be seen with application of Euclidean distance by HRV-Cort as accuracy, sensitivity and specificity scored 80.0%, 83/3%, and 78.3%, respectively. The table also indicated that, biomarkers like PR, BP and HRV representing ANS with Scort had poor performance. From this result, it can be seen that HRV is a strong biomarker and can be utilized for ANS.

Following this, a single biomarker was used to measure the differences between outcomes of new fused feature and stress response classifications. Results indicated that the latter had better performance than single biomarker. However, since SVM model indicated excellent outcomes in categorizing fused biomarker (HRV-SCort), it was then utilized in exhibiting stress response index.

Table 2: Outcomes of SVM classifiers applied in fused biomarker

Fusion Method		Acc %	Sen%	Spe%
Serial	HRV-Cort	63.3	51.7	75.0
	BP-Cort	67.2	68.3	66.7
	PR-Cort	58.3	66.7	48.3
	HRV-BP-PR-Cort	65.0	58.3	71.7
Euclidean distance	HRV-Cort	80.0	83.3	78.3
	BP-Cort	21.7	13.3	28.3
	PR-Cort	59.4	91.7	26.7
	HRV-BP-PR-Cort	68.9	81.7	60.0

*Notes: Acc: accuracy, Sen: sensitivity, Spe: Specificity

3.4. Support Vector Machine Model Fusion

The fused feature vector was then utilized as model with SVM MATLAB toolbox application. The input was placed in SVM model. Values of alpha (α), support vector, scale factor, shift and bias were obtained. The stress response index was recorded with

SVM classification (C_{SV}) algorithm as reference. The algorithm is interpreted as follow:

$$C_{sv} = \sum_i \alpha_i k_l(s_i, x) + b \quad (10)$$

where, si acts as support vector, α is the support vector's weight, b signifying bias and defined as the interception of hyperplane that forms normalized data space by division of two groups, and x is the recorded scale training vector from

$$(\boldsymbol{f}_e)_{scaled} = scale\ factor\ \times (\boldsymbol{f}_{Tr} + shift) \quad (11)$$

Meanwhile k1 is the dot product which is also recognized as linear kernel, as follow:

$$\text{Linear kernel} = \text{dot product} = \text{conj}(s_i).*x = s_iT.*x \quad (12)$$

Hence, based on (13), proposed stress response index (SRi) was indicated as:

$$\text{SR}_i = \sum_i \alpha_i \times (s_i{}^T .* [(\boldsymbol{H}_R) + i(\boldsymbol{C}_R)]_{scaled}) - 0.057 \quad (13)$$

where HRV reactivity occurred via HR and SCort reactivity is via CR.

Figure 4 shows the suggested Stress Response Index (*SRi*) of HRV and salivary cortisol (HRV-SCort) with *ed*-SVM of children's traumatic incidents. The figure indicated that two groups of stress response were divided by index boundary at a value of '0'. Normal stress response was identified with index sample ≥ 0, while other samples exceeding normal stress response were categorized as irregular stress response. Irregular stress response is often made up of participants who had no traumatic experiences during their early years.

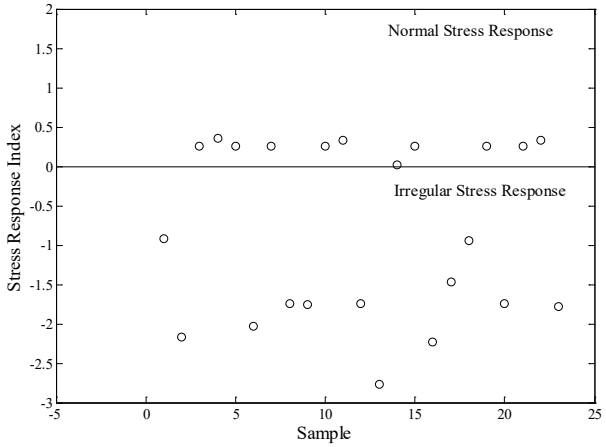


Figure 4: *SRi* of HRV and HRV-SCort with ed-SVM of traumatic childhood experience

A recent stress response index in reference to traumatic childhood incidence was proposed with the obtained comprehensive feature extraction. This extraction is made up of robust HRV feature with the application of GA and integration of *ed* and SVM. Linear regression was referred in designing the stress response index. Hence, the scalar value is inconsistent and changes based on sample's size. Assumption made is that the greater number of sample leads to stress response index being more accurate.

Nevertheless, alcohol intake, smoking, diet, present stress and psychosocial situation may influence the precision of stress response index [64]. This is a take on for researchers to further study on the factors influencing stress response as well as conducting the study with a bigger sample size to create a more complete index that acts as the fundamentals of stress response index. Furthermore, the reaction of traumatic childhood incidents within an individual against the stimulated stress requires further comprehensive research. Chances are that these individuals could be more at risk physically and mentally or they are able to manage themselves being in such situation. Findings from this research can be utilized to upgrade the approach within adult's healthcare [65].

4. Conclusion

This research proposed stress response index based on the combination of biomarker representing both autonomic nervous system and hypothalamus pituitary adrenocortical to classify stress response for traumatic childhood experience. 12 HRV features, form time, frequency, Time, frequency, TF and wavelet were extracted then combined with other biomarkers, blood pressure, pulse rate, and salivary cortisol. The biomarker combination represents two main body systems which are autonomic nervous system and hypothalamus pituitary adrenocortical axis. The classification was performed using support vector machine and the result proved that the combination between HRV and Salivary Cortisol demonstrated the highest performance with 80.0% accuracy, 83.3% sensitivity and 78.3% specificity. Finally, The fused vector was fed into the SVM model to develop the stress response index. This index could be used to predict the future health and therefore allow the individual to take more cautious about his/her health.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research was funded by the Ministry of Higher Education, Malaysian Government (FRGS: R.J130000.7745.4F943) and funding for research from School of Sport, Exercise & Rehabilitation Sciences, University of Birmingham, UK.

References

[1] Aimie-Salleh, Noor & Balakrishnan, Malarvili & Whitttaker, Anna, "Stress response index for adverse childhood experience based on fusion of biomarkers". IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), 140-145, 2018. DOI: 10.1109/IECBES.2018.8626657.

[2] T. D. Washington, Psychological Stress and Anxiety in Middle to Late Childhood and Early Adolescence: Manifestations and Management. Journal of Pediatric Nursing, 24(4), 302-313, 2009.

[3] N. Schneiderman, G. Ironson, and S. D. Siegel, Stress and Health: Psychological, Behavioral, and Biological Determinant. Annual Review of Clinical Psychology, 1, 607–628, 2005.

[4] A. Fernandes, R. Helawar, R. Lokesh, T. Tari, and A. V. Shahapurkar, A Survey on Analysis of Stress. International Journal of Latest Trends in Engineering and Technology, Special Issue, 31-35, 2013.

[5] V. H. Rice, Theories of Stress and Its Relationship to Health. In Rice, V. H. (Ed.) Handbook of Stress, Coping, and Health: Implications for Nursing Research, Theory, and Practice, Detroit: SAGE Publications, Inc, 23-42, 2012.

[6] A. Baum, M. K. O'Keeffe, and L. M. Davidson, Acute Stressors and Chronic Response: The Case of Traumatic Stress. Journal of Applied Social Psychology, 20(20), 1623-1724, 1990.
[7] M. R. Salleh, Life Event, Stress and Illness. The Malaysian Journal of Medical Science: MJMS, Pmc, 15(4), 9-18, 2008.
[8] R. Finlay-Jones, and G. W. Brown, Types of Stressful Life Event and the Onset of Anxiety and Depressive Disorders. Psychol Med, 11(4), 803-15, 1981.
[9] G. Andrews, C. Tennant, D. M. Hewson, and G. E. Vaillant, Life Event Stress, Social Support, Coping Style, and Risk of Psychological Impairment. The Journal of Nervous and Mental Disease, 166(5), 307-16, 1978.
[10] R. F. Anda, D. W. Brown, S. R. Dube, J. D. Bremner, V. J. Felitti, and W. H. Giles, Adverse Childhood Experiences and Chronic Obstructive Pulmonary Disease in Adults. American Journal of Preventive Medicine, 34(5), 396-403, 2008.
[11] F. V. J. Felitti, et al., "Relationship of Childhood Abuse and Household Dysfunction to Many of the Leading Causes of Death in Adults: The Adverse Childhood Experiences (ACE) Study," American Journal of Preventive Medicine, 14, 245-258, 1998.
[12] V. J. Felitti, Long-Term Medical Consequences of Incest, Rape, and Molestation. Southern MedicalJjournal, PubMed 84(3), 328-331, 1991.
[13] V. J. Felitti, Childhood Sexual Abuse, Depression, and Family Dysfunction in Adult Obese Patients: A Case Control Study. Southern Medical Journal, PubMed, 86(7), 732-736, 1993.
[14] W. R. Lovallo, N. H. Farag, K. H. Sorocco, A. J. Cohoon, and A. S. Vincent, "Lifetime adversity leads to blunted stress axis reactivity: studies from the Oklahoma Family Health Patterns Project," Biological Psychiatry, 71, 344-349, 2012.
[15] W. R. Lovallo, "Early life adversity reduces stress reactivity and enhances impulsive behavior: Implications for health behaviors," International Journal of Psychophysiology, 90, 8-16, 2013.
[16] L. J. Luecken, Childhood Attachment and Loss Experiences Affect Adult Cardiovascular and Cortisol Function. Psychosomatic Medicine, 60(6), 765-72. 1998.
[17] L. J. Luecken and D. S. Roubinov, "Hostile behavior links negative childhood family relationships to heart rate reactivity and recovery in young adulthood," International Journal of Psychophysiology, 84, 172-179, 02/11 2012.
[18] R. Anda, D. Williamson, D. Jones, C. Macera, E. Eaker, A.Glassman, and J.Marks, "Depressed Affect, Hopelessness, and the Risk of Ischemic Heart Disease in a Cohort of U.S. Adults," Epidemiology, 4, 285-294, 1993.
[19] R. F. Anda, et al., "The enduring effects of abuse and related adverse experiences in childhood," European Archives of Psychiatry and Clinical Neuroscience, 256, 174-186, 2006
[20] S. V. Batten, D. W. Brown, S. R. Dube, J. D. Bremner, V. J. Felitti, and W. H. Giles, "Childhood Maltreatment as a Risk Factor for Adult Cardiovascular Disease and Depression," Journal of Clinical Psychiatry, 65, 249-54, Feb 2004.
[21] M. Dong, et al., "Insights into causal pathways for ischemic heart disease: adverse childhood experiences study," Circulation, 110, 1761-6, Sep 28 2004.
[22] W. T. Boyce and B. J. Ellis, "Biological sensitivity to context: I. An evolutionary-developmental theory of the origins and functions of stress reactivity," Development and Psychopathology, 17, 271-301, Spring 2005.
[23] J. Ellis, et al., "Biological sensitivity to context: II. Empirical explorations of an evolutionary-developmental theory," Development and Psychopathology, 17, 303-28, Spring 2005.
[24] U. R. Acharya, et al., "Heart Rate Variability: A Review," Medical and Biological Engineering and Computing, 44, 1031-1051, 2006.
[25] A. P. Allen, et al., "Biological and Psychological Markers of Stress in Humans: Focus on the Trier Social Stress Test," Neuroscience and Biobehavioral Reviews 38, 94–124, 2014.
[26] Task Force, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology," Circulation, 93, 1043-1065, 1996.
[27] J. F. Thayer, et al., "A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health," Neuroscience & Biobehavioral Reviews, 36, 747-756, 2012.
[28] J. E. Carroll, et al., "Childhood abuse, parental warmth, and adult multisystem biological risk in the Coronary Artery Risk Development in Young Adults study," Proceedings of the National Academy of Sciences, 110, 17149-17153, 2013.
[29] K. Winzeler, et al., "Adverse childhood experiences and autonomic regulation in response to acute stress: the role of the sympathetic and parasympathetic nervous systems," Anxiety, Stress, & Coping, 30, 145-154, 2017/03/04 2017.
[30] N. Michels, et al., "Relation between salivary cortisol as stress biomarker and dietary pattern in children," Psychoneuroendocrinology, 38, 1512-1520, 2013.
[31] N. Michels, et al., "Children's heart rate variability as stress indicator: Association with reported stress and cortisol," Biological Psychology, 94, 433-440, 2013.
[32] A. T. Ginty, et al., "Disordered eating behaviour is associated with blunted cortisol and cardiovascular reactions to acute psychological stress," Psychoneuroendocrinology, 37,715-724, 5// 2012.
[33] A. C. Phillips, et al., "Cardiovascular and cortisol reactions to acute psychological stress and adiposity: cross-sectional and prospective associations in the Dutch Famine Birth Cohort Study," Psychosomatic Medicine, 74, 699-710, Sep 2012.
[34] D. P. Bernstein, et al., "Development and validation of a brief screening version of the Childhood Trauma Questionnaire," Child Abuse & Neglect, 27, 169-90, Feb 2003.
[35] J. Pan and J. W. Tompkins, "A Real-Time QRS Detection Algorithm," IEEE Transactions on Biomedical Engineering, 32, 230-236, 1985.
[36] U. R. Acharya, et al., "Automatic Identification of Cardiac Health Using Modeling Techniques: A Comparative Study," Information Sciences, 178, 4571-4582, 2008.
[37] R. H. Takalo and H. H. Ihalainen, " Tutorial on Univariate Autoregressive Spectral Analysis Export," The Journal of Clinical Monitoring and Computing, 20, 379-379, 2006.
[38] M. N. Castro, et al., "Heart rate variability response to mental arithmetic stress in patients with schizophrenia: Autonomic response to stress in schizophrenia," Schizophrenia Research, 99, 294-303, 2008.
[39] P. Karthikeyan, et al., "Analysis of Stroop Color Word Test-Based Human Stress Detection using Electrocardiography and Heart Rate Variability Signals," Arabian Journal for Science and Engineering, 39, 1835-1847, 2014.
[40] J. A. J. Heathers, "Smartphone-enabled pulse rate variability: An alternative methodology for the collection of heart rate variability in psychophysiological research," International Journal of Psychophysiology, 89, 297-304, 2013.
[41] B. Barkat and B. Boashash, "A High-Resolution Quadratic Time-Frequency Distribution for Multicomponent Signals Analysis," IEEE Transactions on Signal Processing, 49, 2232-2239, 2001.
[42] N. Aimie-Salleh, "Autonomic Function Assessment Tool Using Time-Frequency Analysis of Heart Rate Variability," Master Thesis, Universiti Teknologi Malaysia, Johor Bahru, 2013.
[43] M. B. Malarvili and M. Mesbah, "Newborn Seizure Detection Based on Heart Rate Variability," IEEE Transactions Biomedical Engineering, 56, 2594-2603, 2009.
[44] D. Giri, et al., "Automated diagnosis of Coronary Artery Disease affected patients using LDA, PCA, ICA and Discrete Wavelet Transform " Knowledge-Based Systems, 37, 274-282, 2013.
[45] G. Kheder, et al., "Heart Rate Variability Analysis Using Threshold of Wavelet Package Coefficients " International Journal on Computer Science and Engineering, 1, 131-136, 2009.
[46] R. Acharya, U., et al., "Computer-aided Diagnosis of Diabetic Subjects by Heart Rate Variability Signals Using Discrete Wavelet Transform Method," Knowledge-Based Systems, 81, 56-64, 2015.
[47] B. Oluleye, A. Leisa, J. Leng, and D. Dean, A Genetic Algorithm-Based Feature Selection. International Journal of Electronics Communication and Computer Engineering, 5(4), 899-905, 2014a.
[48] B. Oluleye, A. Leisa, J. Leng, and D. Dean, Zernike Moments and Genetic Algorithm : Tutorial and Application. British Journal of Mathematics & Computer Science 4(15), 2217-2236, 2014b.
[49] N. Akhter, S. Dabhade, N. Bansod, and K. Kale, Feature Selection for Heart Rate Variability Based Biometric Recognition using Genetic Algorithm. In Berretti, S., Thampi, S. M. & Srivastava, P. R. (Ed.) Intelligent Systems Technologies and Applications, Maharashtra, India: Springer International Publishing, 91-101, 2016.
[50] M. S. Mohamad, S. Deris, S. M. Yatim, and M. R. Othman, Feature Selection Method using Genetic Algorithm for the Classification of Small and High Dimension Data. First International Symposium on Information and Communications Technologies. 7-8 October. Putrajaya, Malaysia, 1-4, 2004.
[51] H. Shahamat, and A. A. Pouyan, Feature Selection using Genetic Algorithm for Classification of Schizophrenia using FMRI Data. Journal of AI and Data Mining, 3(1), 30-37, 2015.
[52] S. Ollander, "Wearable Sensor Data Fusion for Human Stress Estimation," Master Master Thesis, Department of Electrical Engineering, Linköping University, Linköping, 2015.
[53] J. Yang, et al., "Feature fusion: parallel strategy vs serial strategy," 36, 1369-1381, 2003.
[54] N. Sharma and T. Gedeon, "Objective measures, sensors and computational techniques for stress recognition and classification: A survey," Computer Methods and Programs in Biomedicine, 108, 1287-1302, 2012.
[55] J. Zhai and A. Barreto, "Stress Detection in Computer Users Based on Digital Signal Processing of Noninvasive Physiological Variables," in Annual International Conference of the IEEE Engineering in Medicine and Biology Society, New York, 2006, 1355-1358.

[56] Z. B. Siao, et al., "Human emotional stress analysis through time domain electromyogram features," in 2013 IEEE Symposium on Industrial Electronics & Applications, 2013, 172-177.
[57] J.-S. Wang, et al., "A k-nearest-neighbor classifier with heart rate variability feature-based transformation algorithm for driving stress recognition," Neurocomputing, 116, 136-143, 2013.
[58] A. Birkhofer, et al., "The deceleration capacity – a new measure of heart rate variability evaluated in patients with schizophrenia and antipsychotic treatment," European Psychiatry, 28,81-86, 2013.
[59] A. Deepak, et al., "Time Domain Measures of Heart Rate Variability during Acute Mental Stress in Type 2 Diabetics: A Case Control Study," National Journal of Physiology, Pharmacy & Pharmacology, 4, 34 – 38, 2014.
[60] A. Muaremi, et al., "Towards Measuring Stress with Smartphones and Wearable Devices During Workday and Sleep," BioNanoScience, 3, 172-183, 2013.
[61] R. Bailon, et al., "Time-Frequency Analysis of Heart Rate Variability During Stress Testing Using "A Priori" Information of Respiratory Frequency," in Computers in Cardiology, Valencia, Spain, 169-172, 2006.
[62] H. M. Seong, et al., "The Analysis of Mental Stress using Time-Frequency Distribution of Heart Rate Variability Signal," in Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, California, 2004, 283-295.
[63] W. S. Liew, et al., "Classifying Stress From Heart Rate Variability Using Salivary Biomarkers as Reference," IEEE Transactions on Neural Networks and Learning Systems, 27, 2035-2046, 2016.
[64] K. A. Kendall-Tackett, "The health effects of childhood abuse: four pathways by which abuse can influence health," Child Abuse & Neglect 26, 715-29, Jun 2002.
[65] J. P. Shonkoff, et al., "Neuroscience, molecular biology, and the childhood roots of health disparities: Building a new framework for health promotion and disease prevention," JAMA, 301, 2252-2259, 2009.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 320-327 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Methodology for Assessing Synchronization Conditions in Telecommunication Devices

Juliy Boiko[*,1], Ilya Pyatin[1], Oleksander Eromenko[2], Oleg Barabash[3]

[1]*Khmelnytskyi National University, Department of Telecommunications and Radio Engineering, 29016, Ukraine*

[2]*Khmelnytskyi National University, Department of Physics and Electrical Engineering, 29016, Ukraine*

[3]*State University of Telecommunications, Department of Mathematics, 03110, Ukraine*

ARTICLE INFO

Article history:
Received: 11 January, 2020
Accepted: 10 March, 2020
Online: 23 March, 2020

Keywords:
Synchronization
Modulation
Dispersion of Error
Interpolation
Timing Error
Signal to Noise Ratio

ABSTRACT

This paper represents analysis of principles providing maximally reliable delay estimation of signal synchronization devices in telecommunication. Clock synchronization schemes on the base of stochastic models, graph-analytic interpretation, analytical functional description and their parameters selection in case of real complex hindrance are analyzed. Dependences between dispersion of maximally reliable estimation of synchronization devices and signal to noise ratio (SNR) are represented. Analytical expressions and dependences are obtained for deviation of synchronization error with respect to SNR at the output of discriminator. Diagrams of state are formed for assessment of digital synchronization device in case of error state. Calculated error of numerical solution of the equation of a stationary grid method. We obtain the schedule of the distribution of the probability density synchronization errors for different values of the normalized time. The functions performed by the symbolic synchronization circle, features of the construction of the timing error detector (TED) are analyzed. The Gardner, Early-Late, Mueller & Muller TED, their error and the S-curve equation are investigated. The dependence of the relative timing error on the SNR is obtained. The article examines the effects of interpolation and decimation conditions in the forming filters of digital communication systems.

1. Introduction

This paper is an extension of work originally presented in conference name "Quality Assessment of Synchronization Devices in Telecommunication" that is published in 2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO) [1].

The proposed paper (additionally) presents experimental results for evaluating the effect of interpolation conditions and decimation for energy gain of telecommunication facilities on phase shift keying modulation, as well as comparative analysis of error detectors of symbol synchronization of coherent communication network with QPSK modulation has done.

The main purpose of the presented results is to supplement and summarize information on the specifics and nature of synchronization conditions in the means of determining the factors that qualitatively influence the factors contributing to the noise immunity of modern information processing systems [2,3]. The coherent and quasicoherent methods for signal receiving require phase locked loop (PLL) frequency as obligate component which forms the reference signal from the received oscillation [4]–[6]. To receive discrete information a system needs clock, word and frame synchronization. The important task of synchronization, which is to form reference oscillation at the receiving side of wireless connection, in most of the cases, becomes complicated due to influence of noise, which distorts the received signal, and in some cases due to random character of the signal itself [2].

Such events entail fluctuational phase deviation, which is formed by reference oscillation synchronization system. The effectiveness of data transmission lowers consequently because of losses in transmission of data required for not only synchronization but also for the signal search time and transition onto monitoring mode for all synchronization systems of the receiver. This raises the question of improving synchronization system especially in

[*]Corresponding Author : Juliy Boiko, Khmelnytskyi National University, 29016 Ukraine, boiko_julius@ukr.net

case of using in communication systems frequency, phase and quadrature methods of digital modulation [7]–[9].

Papers [4,5] represent linear and nonlinear analysis of clock synchronization devices. Probability of acquisition and probability of cycle slip are estimated. The main factors, which increase an error in increasing and decreasing reliability of system synchronization in time, are analyzed.

The represented research is conducted in progress of results published in [3]–[6]. We use various stochastic models with expanded technical interpretation of Markov's modeling, cumulant method and method for statistic linearization. Besides, upgrading circuit structure of the manipulated signal receiver in order to optimize it by using solutions aimed at removing reverse effect, synthesis of the receiver structure with synchronization devices and defining character of signal to noise ratio influence on accuracy of synthesized clock synchronization devices are important.

The research outcome represents novelty, which consists in using stochastic model to describe synchronization error. Significance of the research consists in estimating clock transition distribution density. The criteria and algorithm is developed to provide optimal operation of synchronizer in condition of hindrance.

2. Stochastic Model of Synchronization System

2.1. Semigraphical Interpretation of Clock Synchronization Systems

In article [1], [7], a stochastic model of a synchronization system was proposed. Model of synchronization system is described by stochastic equations in accordance to the scheme Fig. 1. The following symbols are used: $\xi=\xi(t)$ - input influence, $\hat{\xi}=\hat{\xi}(t)$ - influence assessment, $\gamma=\gamma(t)=(\xi(t)-\hat{\xi}(t))/T$ - normalized synchronization error, $\rho(\gamma)$ - discriminatory characteristics, $h(\gamma)$ - fluctuation characteristics, $n(t)$ - Gaussian white noise, $f(\varepsilon)$ - transient pulse function of the linear dynamic element, which describes the effect of processing output signal and adjustment of the clock generator frequency Fig.1.

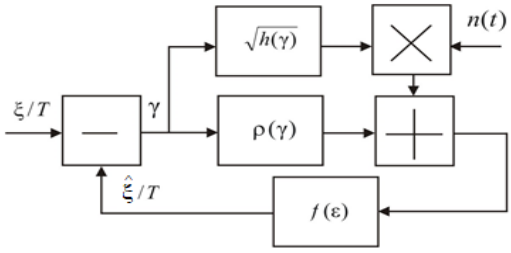


Figure 1: The research clock synchronization scheme

Assumption is made that: the influence is slow process, stable due to large amount of symbol intervals; assessment of the target influence is a slow process; normalizing random processes and fluctuation of discriminator's output signal is the condition to be used; spectral density $G(\omega, \gamma)$ is assumed stable in the tract transmission frequency band $G(0, \gamma)$ and fluctuation component is represented as the white noise.

The synchronization system is described as automatic control system. The input influence is represented as $r(t)=U(t)+n(t)$. The signal is represented analogously as the following: $U(t)=U_0\sum_i S_i\, h(t-iT-\xi)$, $h(t)=1$, $0\le t\le T$, $h(t)=0$, $t\in(0,T)$.

Thus, the pulses sequence is given U_0 - having rectangular form $h(t)$, which represents binary character data stream $S_i\in\{+1,-1\}$, $n(t)$ - Gaussian white noise. The research is conducted in order to identify discriminatory and fluctuational characteristics. The synchronization error is specified by $\gamma=(\xi-\hat{\xi})/T$. The pulse sequences to identify synchronization error are represented in Fig. 2.

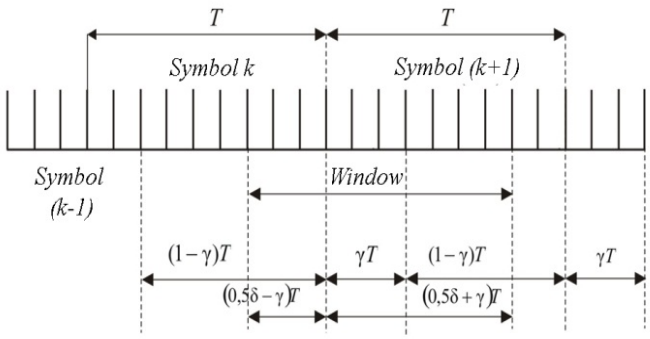


Figure 2: Pulse sequences in synchronization system

The dispersion of fluctuational error in synchronization system is to be defined using the equation of Markov's random process, which density of probability distribution $P=P(\gamma, t)$ is expressed by Fokker-Planck equation [10,11]:

$$\frac{d\gamma}{dt}+k\alpha(\gamma)=k\sqrt{h(\gamma)}n(t),\quad 0\le t, \tag{1}$$

$$\gamma(0)\sim P_0(\gamma).$$

where $P_0(\gamma)$ - initial density distribution of synchronization error.

$$\frac{\partial P}{\partial t}=-\frac{\partial}{\partial\gamma}\left(A(\gamma)P\right)+\frac{1}{2}\frac{\partial^2}{\partial\gamma^2}\left(Z(\gamma)P\right),\quad P(\gamma,0)=P_0(\gamma) \tag{2}$$

where $U_0(\gamma)=-k\alpha(\gamma)$, $Z(\gamma)=k^2S(0,\gamma)$.

The solution of equation (2) at synchronization system stationary mode $\frac{\partial P}{\partial t}=0$ is found as:

$$P(\gamma)=c^{-1}\exp\left(-\int_0^{\gamma}\frac{2h_0^2\chi_0\alpha_n(U)+dh(U)/dU}{h(U)}dU\right),\quad |\gamma|\le 1/2 \tag{3}$$

where $\chi_0=4/U_0kT$ - parameter opposite to normalized noise band of the synchronization system linear model [12]-[15]; $c=2\int_0^{1/2}\exp\left(-\int_0^{\gamma}\frac{2h_0^2\chi_0\alpha_n(U)+dh(U)/dU}{h(U)}dU\right)d\gamma$ - normalizing constant.

Synchronization system accuracy is assessed by dispersion of normalized synchronization error:

$$\sigma_\gamma^2 = 2\int_0^{1/2}\gamma^2 P(\gamma)d\gamma \qquad (4)$$

The mathematical modeling will be performed and error dispersion influence will be assessed in dependence on signal to noise ratio for some values of χ_0 parameter. In particular Fig. 3 demonstrates assessment carried out for $\delta_0=\delta_{0opt}$ (h_0^2).

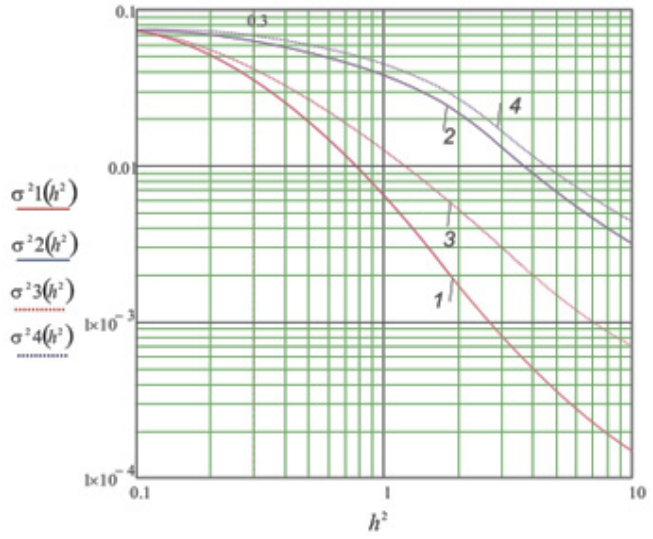


Figure 3: Graphs of dispersion error σ_γ^2 dependence on signal to noise ratio h_0^2: 1- $\chi_0=100, \delta_0\rightarrow$opt; 2- $\chi_0=25, \delta_0\rightarrow$opt; 3- $\chi_0=100, \delta_{0=}1$; $\chi_0=25, \delta_0=1$

The conducted research has indicated that in case of narrowing synchronization system band the dispersion of fluctuational error reduces. In Fig. 3 solid lines 1 and 2 demonstrate effect that the window width minimizes synchronization error dispersion for each value of signal to noise ratio, however lines 3 and 4 demonstrate constant width of the window.

Analysis of transition mode of synchronization system can be done with the use of numerical solving Fokker-Planck equation by grid method when distribution density of synchronization error changes. The equation (5) will be used with inserting normalized time $t_H=4\varpi t$, where $\varpi=k_1k_2U_0T/4$ - noise band of linear model of synchronization system. In this case the canonic form of stochastic Fokker-Planck equation for distribution density $P=P(\gamma, t_H)$ can be written.

$$\frac{\partial P}{\partial t_H}=\frac{\partial}{\partial\gamma}(\alpha(\gamma)P)+\frac{1}{D_0}\frac{\partial^2}{\partial\gamma^2}(h(\gamma)P), \qquad (5)$$

$$t_H\geq 0, \quad P(\gamma,t_H)=P(\gamma\pm1,t_H), \quad P(\gamma,0)=1$$

In the research the distribution density of the initial condition will be assumed uniform, $D_0=2h_0^2/\varpi T$ - equation parameter, $\Delta F_{nor}=\varpi T$ - normalized noise band.

Stationary solution of equation (5) is used for numerical solution (2) by grid method. The essence of this method is in inserting the grid having nodes $x_k=x_0+k\xi$, $k=0,\pm1, \pm2,\ldots,\pm N$, $x_0=0$, $\xi=1/2N$, $t_l=t_0+\delta_l$, $l=0,1,2,\ldots$, $\tau_0=0$, ξ, δ - grid. The grid function is represented as $\hbar_{k,l}=\hbar(\gamma, t_H)$ when $\gamma=x_k$, $t_H=t_l$ and differential operators are replaced with difference ones. Initial conditions are: $P_{k,0}=1$, $k\in(0, \pm N)$, $P_{-N,l}=P_{N,l}$, $P_{N+1,l}=P_{-N+1,l}$, $P_{-N-1,l}=P_{N-1,l}$. References testify that solution is possible to find at the new layer on conditions mentioned above [16]. The stability condition of the difference scheme is worthy to mention: $\delta/\xi^2 < D_0/2$.

The equation (5) rewritten with recurrent formulas becomes:

$$P_{k,l+1}=P_{k,l}+\frac{\delta}{2\xi}\left(\alpha_{k+1}P_{k+1,l}-\alpha_{k-1}P_{k-1,l}\right)+$$

$$+\frac{\delta}{D_0\xi^2}\left(h_{k+1}P_{k+1,l}-2h_kP_{k,l}+h_{k-1}P_{k-1,l}\right) \qquad (6)$$

$l=0,1,2,\ldots$ $\omega(\varphi,\tau)=1/2\pi\cdot(ch\tau-ch\tau\cdot\cos\varphi)$.

The error should be assessed by comparing grid functions of two types – the accurate and the numerical solutions; and by finding $\Delta=(\omega_a(\varphi,\tau)-\omega_n(\varphi,\tau)/\omega_a(\varphi,\tau))\cdot100\%$ along the layers in nodes of selected grid.

This approach will be used to assess error of clock synchronization. Fig. 4 shows transient process, Fig. 5 – dependences of numerical solution (6), and Fig. 6 represents graphs, which describe process of setting on stationary mode at error probability $p_0=10^{-3}$, in case of minimum shift keying (GMSK) digital modulation [1].

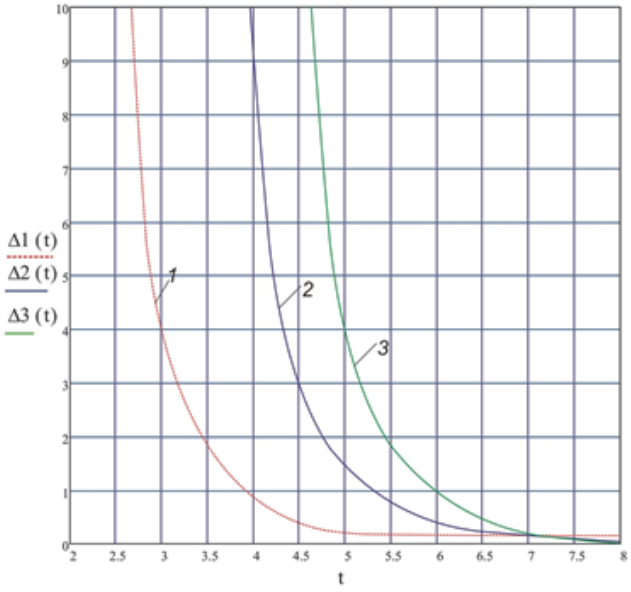


Figure 4: Graphs of transient processes t_H (condition: $p_0=10^{-3}$); 1- $\varpi t=5\cdot10^{-2}$; 2- $\varpi t=1\cdot10^{-2}$; 3- $\varpi t=1\cdot10^{-3}$

Modeling parameters are: $\Delta F_{nor}=5\cdot10^{-2}$; $\xi=0.01$, $\delta=0.001$.

The research has identified less than one percent error in numerical search of solution using grid method Fig. 5.

The research results in form of graphs which describe transient process for the value of SNR that corresponds to error probability $p_0=10^{-3}$ (Fig. 6) and assessed dependences of transient process duration in a system in case of deviation of current dispersion $\sigma^2_{current}$ of synchronization error on its value in stationary mode $\sigma^2_{stationary}$, testify that in case of narrowing band in synchronization

system, what means reducing monitoring error, the transient process time gets increased (Fig. 4).

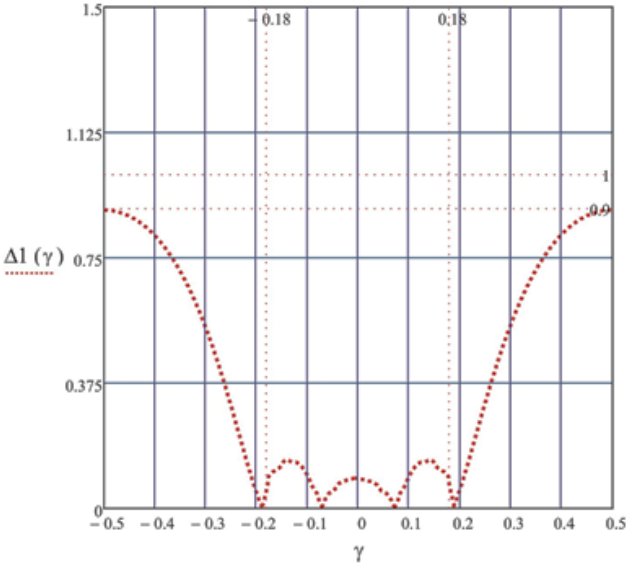


Figure 5: Numerical stationary solutions by grid method: Δ - given in percent's

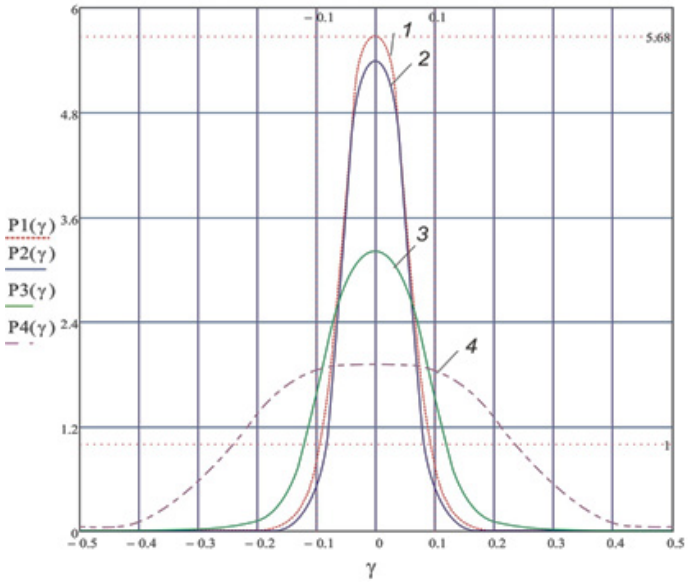


Figure 6: Graphs of density of probability distribution at different values of normalized time t_H (condition: $p_0=10^{-3}$, $\varpi t=5\cdot10^{-2}$): 1 – stationary state; 2 - $t_H=2$; 3 - $t_H=1$; 4 - $t_H=0.5$

Besides, in case when the band of synchronization system changes then getting on stationary mode goes with larger number of symbols.

3. Investigation Inaccuracy of Symbol Timing Error Detector

In the modern digital communication system, the sampling rate plays an important role. High-speed signal processing can be implemented by usage of digital converters with increasing or decreasing sampling rate [17].

An important part of the digital interface used for radio frequency communication systems is the frequency converter (FC). The FC main function is to convert one or more data channels from the main band format into a radio signal consisting of modulated carriers belonging to a set of one or more mentioned radio frequencies. This is achieved in two stages: increasing the sampling rate by interpolation, providing spectrum generation and suppressing interpolation images by filtration and shifting the spectrum of the signal to the desired carrier frequencies using a multiplier and a heterodyne [18,19].

3.1. Symbol Timing Recovery Circuit

The symbol timing recovery circuit has the block diagram shown in Fig. 7. Time synchronization is performed using a PLL consisting of three main units: TED, contour filter and interpolation control [20, 21].

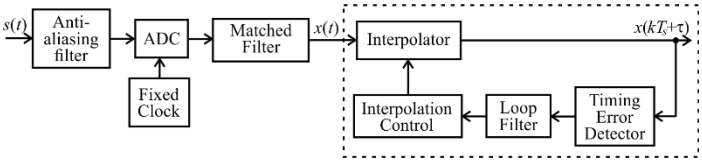


Figure 7: Structural diagram of the symbolic synchronization circle

The figure $x(t)$ shows signal sampling after the matched filter; $x(kT_S+\tau)$ is a symbolic signal adjusted for the offset of the clock signal between the transmitter and receiver.

The purpose of characters synchronization is to create N samples at the output of the matched filter for each character interval so that one of the samples is aligned with the maximum opening of the eye-diagram [7]. A sample rate of the received signal $u(t)$ is received at a fixed rate $1/T$, which is asynchronous with the rate of symbols transmission. The ADC generates samples of the input signal at a rate of N samples per symbol. The N-th count of this signal can be represented by the expression:

$$r(nT) = G_a \sum_m a(m)\, p(nT - mT_s - \tau) + w(nT)$$

The time delay τ is estimated by samplings $x(nT)$ which are asynchronous samplings at the output of the matched filter. Samples are not aligned in characters. The role of symbol synchronization over time is to "move" samples to the right moments. Another name for "moving" samplings over time is interpolation [20]. Because the time synchronizer must adapt to an unknown time delay, the interpolator must be adaptive. With proper operation, the interpolator generates agreed output filter signals that are aligned across characters and optimal sampling points.

The main disadvantage of the scheme shown in Fig. 7, is an interpolation jitter that occurs when $T_I \neq NT$. In this case the result of the interpolation is displayed in average for each N sample. But on condition of $T_I \neq NT$, the fractional timing error accumulates and as a result it reaches the value of the character length. At a time when this is happening the interpolator outputs N-1 of the samples or N+1 after the previous interpolated samples to compensate for the difference depending on the sign of the fractional timing error. This interpolation jitter is particularly problematic if the data bits need to be retransmitted synchronously to some other destination.

In general, the TED generates an error signal once for each character based on the current synchronization estimate and using

the matched filter input $r(nT)$ and the matched filter output $x(nT)$. In other words, the discrete-time error signal is updated with a frequency of characters.

It can be supposed that there is an ideal interpolator that calculates the interpolation polynomial $x(kT_s+\hat{\tau})$ using the estimation $\hat{\tau}$ of the time delay and the output signals of the matched filter. The interpolation polynomial can be represented by the expression:

$$x(kT_s+\hat{\tau}) = K\sum_m a(m)r_p((k-m)T_s-\tau_e)+v(kT_s+\hat{\tau}) \quad (8)$$

where $\tau_e=\tau-\hat{\tau}$ - is a timing error and $K=G_a/T$. The TED generates a signal that is a function of the timing error τ_e. The TED output signal $e(kT_s)$ is a function of the interpolated output signals of the matched filter and the data symbols (or their estimates). The TED characteristics are described by the *S*-curve $g(\tau_e)$.

Symbolic synchronization relies on a change of data sign when determining a synchronization error signal. For a QPSK modulation communication system the error signal oriented to symbol-synchronization system solution is determined by the expression:

$$e(k) = sign\{a_0(k)\}\dot{x}(kT_s)+sign\{a_1(k)\}\dot{y}(kT_s) \quad (9)$$

where $\dot{x}(kT_s)$ and $\dot{y}(kT_s)$ - are derivatives in time outputs of respectively common-mode and quadrature matched filters.

3.2. Error Signal Characteristics

The derivative of the correlation function of the matched filter is shown in Fig. 8.

This error signal is derived from an estimate of the maximum likelihood of synchronization bias. The operation of a derivative obtaining can be replaced by the structure of management of early-late counts for both common-mode and quadrature components.

Character synchronization according to Gardner is a non-data feedback method. The timing error detector requires at least two samples per symbol, one of which is the point where a decision can be made.

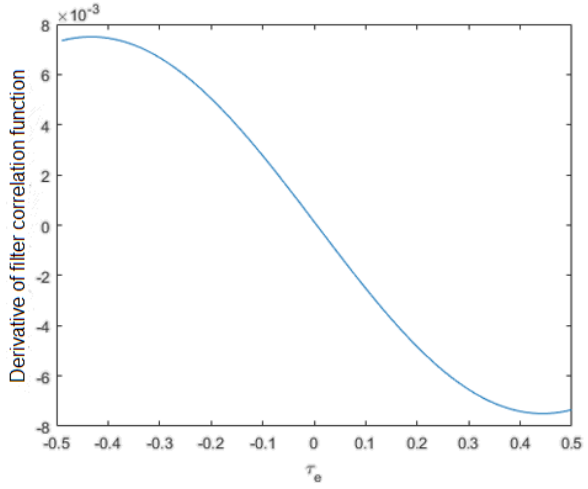


Figure 8: Derivative of autocorrelation function of spectral raised-cosine pulse type filter

The Gardner timing error detector is based on the detection of zero crossing on eye diagrams and is designed to work with BPSK and QPSK [7].

For a Gardner detector at arbitrary fixed $\hat{\tau}$, the synchronization error $e(k)$ is independent of any rotation of the carrier phase [17]. The Gardner TED has an error signal defined by the expression:

$$e(k) = x((k-1/2)T_s+\hat{\tau})\left[x((k-1)T_s+\hat{\tau})-x(kT_s+\hat{\tau})\right] \quad (10)$$

S-curve for GTED turns out by transformation $e(k)$ through τ_e and calculating the desired value:

$$g(\tau_e) = K^2E_{avg}\sum_m r_p((m-1/2)T_s-\tau_e)\left[r_p((m-1)T_s-\tau_e)-r_p(mT_s-\tau_e)\right]$$

The Early-Late detector restores the phase of input character synchronization using a non-data feedback method.

The optimum moment of sampling signals coincides with the moment of maximum average eye-diagram opening. The maximum eye opening diagram occurs at a time when the average slope of the eye-diagram is zero. The nonzero slope at $t=\tau$ - are the points on the trajectory corresponding to the absence of the data sign change. Symbolic synchronization relies on data sign changes to get the correct synchronization error signal. The TED Early-Late works with 2 samples per symbol. The data-dependent detector synchronization error signal looks like:

$$e(k) = a(k)\left[x((k+1/2)T_s+\hat{\tau})-x((k-1/2)T_s+\hat{\tau})\right]$$

where T_s - is usually chosen as a value that is conveniently represented by the sampling rate.

The *S*-curve for TED Early-Late is obtained by calculating the expected value of the error signal:

$$g(\tau_e) = KE_{avg}\left[r_p(T_s/2-\tau_e)-r_p(-T_s/2-\tau_e)\right] \quad (12)$$

Shedule of *S*-curve is shown in Fig. 9.

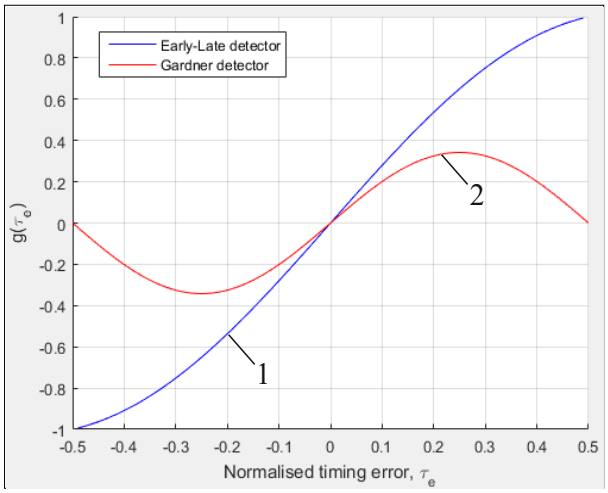


Figure 9: *S*-curve of detectors of timing error (1 – is a Early-Late detector; 2 – is a Gardner detector)

The Mueller & Muller detector implements a decision-oriented method of data-driven feedback that requires a preliminary restoration of the carrier phase [22]. The Mueller & Muller timing error detector operates on matched filter output signals generated at 1 sample per character. The symbol synchronization error signal is determined by the expression:

$$e(k) = a(k-1)x(kT_s + \hat{\tau}) + a(k)x((k-1)T_s + \hat{\tau}) \quad (13)$$

where $\hat{\tau}$ - is the estimation of clock delay; $x(kT_s + \hat{\tau})$ - is the interpolation polynom; $a(k)$ - is the k-th received signal symbol; T_s - is the duration of the character; τ – is the unknown time delay.

The S-curve is obtained by calculating the expected value $e(k)$ for $x(kT_s-\tau_e)$ and $x((k-1)T_s-\tau_e$. The S-curve for the detector of data-dependent synchronization error is determined by the expression:

$$g(\tau_e) = KE_{avg}\left[r_p(T_s - \tau_e) - r_p(-T_s - \tau_e)\right] \quad (14)$$

The S-curve is the slope estimation of $r_p(\tau_e)$ using the values $r_p(t)$ of time symbol before and after $-\tau_e$. Since $r_p(t)$ is an autocorrelation function, it is symmetric about τ_e=0. As a result, the S-curve is zero at τ_e=0.

The dependence schedule of the relative synchronization error on the SNR for the TED of Early-Late, Gardner and Mueller & Muller is shown in Fig. 10.

So, the relative synchronization error increases when the SNR is less than 10 dB. Early-Late TED has the most uniform dependence.

When demodulating digital signals, it is desirable that the sampling rate of the signal can be multiple of its keying rate (each character must have the same number of counting's). However, the sampling rate of the input signal from the ADC is generally fixed but the manipulation speed may change. The solution to this problem is the signal resampling.

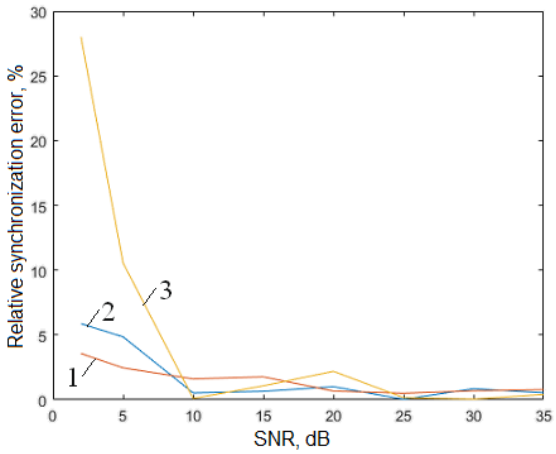


Figure 10: Dependency of relative synchronization error on the SNR for TED (1 – is the Early-Late detector; 2 – is the Gardner detector; 3 – is the Mueller & Muller detector)

4. Conditions of Interpolation and Decimation Signals

The phase modulation techniques should be used to study the effect of conditions of interpolation and decimation signals on telecommunication facilities power.

The estimation should be carried out by determining the energy gain in case of change of the coefficients of interpolation and decimation of the agreed forming filters. Recommendations should be given as to the conditions for changing the coefficients and using the symbolic synchronization system to stabilize the modes of the digital communication system when using interpolation and decimation in the forming filters.

Background information for estimating the effect of interpolation and decimation signals on the power of telecommunications facilities with phase manipulation will be presented by noise immunity graphs (BER dependence on E_b/N_0) and evaluation of signal processing quality by constellation diagrams (signal constellations: Scatter Plot) [7]. It is necessary to evaluate the dependencies of the bit error rate (BER) on the SNR (E_b/N_0) for the different graphs of error synchronization detector and interpolation (decimation) coefficients in the forming filters.

Fig. 11, Fig.12 and Fig.13 show the results of the study of the bit error rate (BER) on the SNR (E_b/N_0) at the coefficients of interpolation and decimation (k_i-d) of matched forming filters equal to 10 and 50 for different types of modulation. Simulink models of digital systems with different types of modulation and different coefficients of interpolation and decimation (k_i-d) of forming filters for QPSK modulation are investigated.

To achieve BER=10^{-4} when using QPSK modulation and increasing the interpolation coefficient - decimation of forming filters from 2 to 50, it is necessary to increase the ratio (E_b/N_0) by 14 dB. The most energy-efficient is BPSK modulation. In this case, increasing the coefficients of interpolation and decimation from 10 to 50 requires an increase in the ratio (E_b/N_0) by 7 dB for BER=10^{-5}.

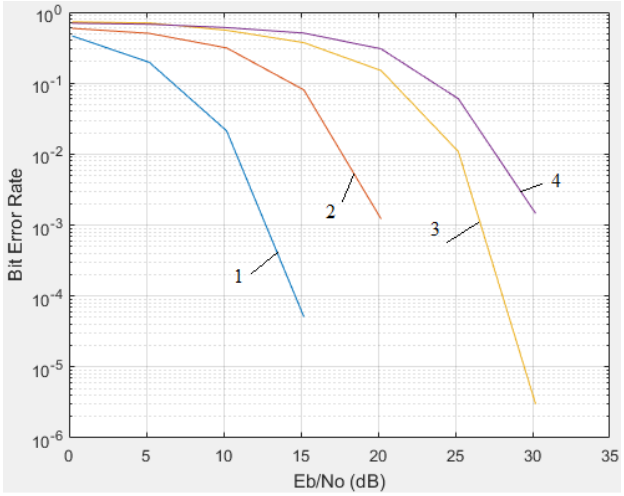


Figure 11: Noise immunity curves of communication system for QPSK modulation with interpolation and decimation coefficients (k_i-d) of forming filters (1 – k_i-d=2; 2 – k_i-d =10; 3 – k_i-d =50; 4 – k_i-d =100)

The timing error detector produces an error signal e, which is related to the difference between the unknown time offset τ and the offset estimate in time space $\hat{\tau}$. The PLL adjusts the estimate $\hat{\tau}$ so that it is close to the time offset τ and reduces the error signal to zero. The logarithmic likelihood ratio L is differentiated with τ to get the error signal and the equation $dL/d\tau_e$ which is solved by PLL.

The Gardner detector and the Mueller-Muller detector are the most commonly used as a synchronization error detector.

The dependence of the bit error rate (BER) on the SNR (E_b/N_0) for different timing error detector circuits and the interpolation and decimation coefficients 2 and 50 in the forming filters for QPSK modulation are shown in Fig.14 and Fig. 15.

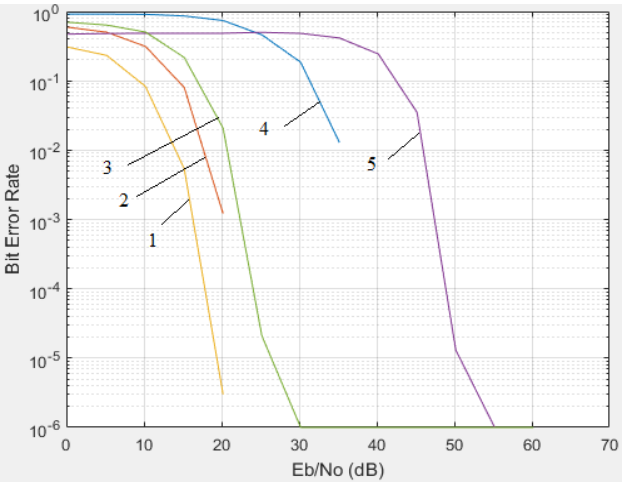


Figure 12: Noise immunity curves (Bit error rate (BER) on the SNR (E_b/N_0) at k_i-d=10 for modulation (1 – BPSK; 2 – QPSK; 3 – DQPSK; 4 – QAM-16; 5 - DBPSK)

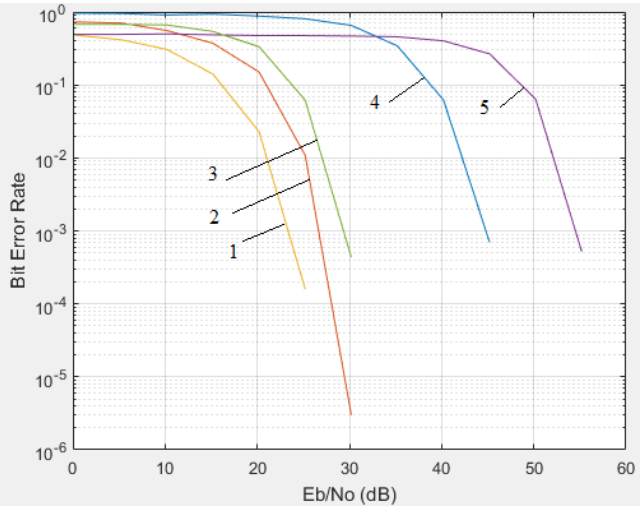


Figure 13: Noise immunity curves (Bit Error Rate (BER) on the SNR (E_b/N_0) at k_i-d =10 for different types of modulation (1 – BPSK; 2 – QPSK; 3 – DQPSK; 4 – QAM-16; 5 - DBPSK)

The Gardner detector has feedback that is unrelated to data and it is independent of carrier frequency recovery [23]. It is suitable for both basic bandwidth and modulated carrier systems. In the presence of noise, the effect of the resynchronizing method improves as the excess bandwidth increases (the recession rate for a spectral raised-cosine pulse type filter).

The Mueller-Muller detector has decision-making feedback that requires a prior restoration of the carrier phase. This method has no noise of its own when the input signal has Nyquist pulses (in the output of a spectral raised-cosine pulse type filter).

Fig. 16 a, b (Scatter Plot1, Scatter Plot2) presents a constellation diagram showing the signal constellation before and after the symbolic synchronization circle and showing the effect of the synchronization error on the transmitted constellation [7], [24].

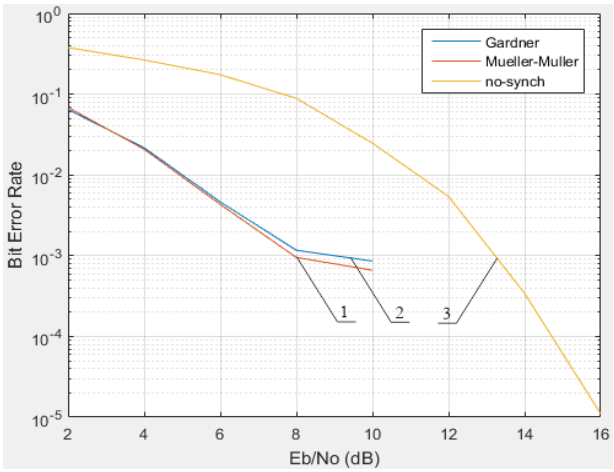


Figure 14: Noise immunity curves of communication system with QPSK modulation and k_i-d = 2 (1 – is a Mueller & Muller detector; 2 – is a Gardner detector; 3 – is the absence of symbol synchronization)

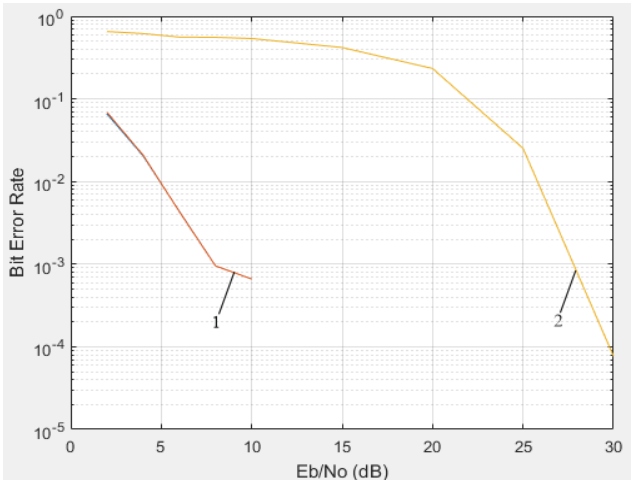
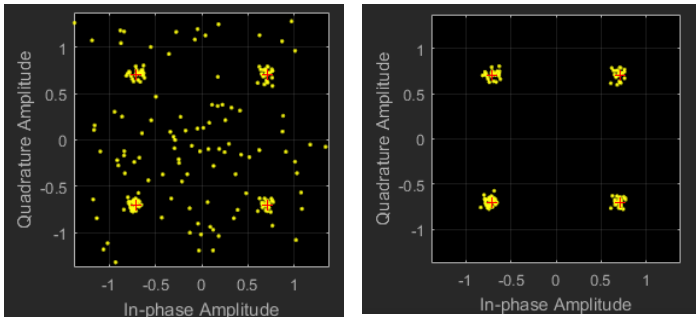


Figure 15: Noise immunity curves of communication system with QPSK modulation and k_i-d=50 (1 – is a Gardner detector; 2 – is the absence of symbol synchronization)



a b

Figure 16: Constellation diagrams (a - Scatter Plot1; b - Scatter Plot2)

In the presence of noise, the output of the Mueller-Muller method improves as the excess pulse width ratio decreases and it is good for narrowband communication systems. The symbol synchronizer unit corrects the asymmetry of the clock pulses and restores the frame synchronization of the QPSK system, which has

a variable synchronization error [25,26]. For a signal to noise ratio of 20 dB, the resulting bit error rate is less than 10^{-3}, indicating that the signal is being restored.

5. Conclusions

Paper represents obtained analytical and graphical research that bonds influence of SNR and accuracy of clock synchronization device on condition of their equal noise bands defined that at the level of $\sigma^2_\gamma=10^{-3}$. The benefit proposed methodology for of clock synchronization device makes almost 2.7 dB.

Studies of Mueller & Muller, Gardner, Early-Late timing error detectors have been performed. As the SNR decreases, there is an increase of the synchronization error. The Mueller & Muller detector has a 20% greater synchronization error than the Gardner and Early-Late detectors with a 2 dB SNR. The Early-Late detector has a larger S-curve, the most uniform dependence of the relative synchronization error on the SNR, which decreases from 4% to 1% with increasing SNR from 2 dB to 35 dB.

To achieve BER $=10^{-4}$ when using QPSK modulation and increasing the interpolation-decimation coefficient of forming filters from 2 to 50, it is necessary to increase the ratio (E_b/N_0) for 14 dB. The most energy-efficient is BPSK modulation. In this case, increasing the coefficients of interpolation and decimation from 10 to 50 requires an increase in the ratio (E_b/N_0) by 7 dB for BER=10^{-5}.

To achieve the bit error rate BER=10^{-3} when using symbolic synchronization, it is possible to reduce the SNR of synchronization for the interpolation and decimation coefficient of the forming filters equal to 2. When the interpolation and decimation coefficient of forming filters increases up to 50, the SNR at the receiver input reduces to 20 dB. Thus, the introduction of a symbolic synchronization system helps to stabilize the operation of the digital communication system when using interpolation and decimation in forming filters.

References

[1] J. Boiko, O. Eromenko, I. Kovtun, S. Petrashchuk, "Quality Assessment of Synchronization Devices in Telecommunication," in 2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO), Kyiv, Ukraine, 2019. https://doi.org/10.1109/elnano.2019.8783438.

[2] A. I. Semenko, N. I. Bokla, M. Y. Kushnir, G. V. Kosovan, "Features of creating based on chaos pseudo-random sequences," in 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), Slavske, Ukraine, 2018. https://doi.org/10.1109/tcset.2018.8336383.

[3] W. Wei, Z. Hongen, C. Suzhi, "Error Analysis and Correction of Time-Frequency Synchronization Based on Packet Switching," in 2017 International Conference on Dependable Systems and Their Applications (DSA), Beijing, China, 2017. https://doi.org/10.1109/dsa.2017.40.

[4] D.R. Stephens, Phase-Locked Loops for Wireless Communications, Springer, 2002.

[5] F. Brandonisio, M.P. Kennedy, Noise-Shaping All-Digital Phase-Locked Loops, Springer, 2014.

[6] S. Huang, S. Liu, M. Liu, J. Hu, Z. Zhu, "Low-Noise Fractional-N PLL With a High-Precision Phase Control in the Phase Synchronization of Multichips," IEEE Microwave and Wireless Components Letters, 28(8), 702-704, 2018. https://doi.org/10.1109/lmwc.2018.2842680.

[7] J. Boiko, V. Tolubko, O. Barabash, O. Eromenko, Y. Havrylko, "Signal processing with frequency and phase shift keying modulation in telecommunications," Telkomnika (Telecommunication Computing Electronics and Control), 17(4), 2025-2038, 2019. http://dx.doi.org/10.12928/telkomnika.v17i4.12168.

[8] A. Rai, V. N. Kumar, "Wideband acquisition technique for QPSK demodulator," in 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 2016. https://doi.org/10.1109/rteict.2016.7807869.

[9] L. Zhang, Z. He, "A modified timing synchronization algorithm for QPSK in digital receiver," in 2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), Dengleng, China, 2011. https://doi.org/10.1109/aimsec.2011.6011207.

[10] P. Bondalapati, W. Namgoong, "Timing Jitter Distribution and Power Spectral Density of a Second-Order Bang–Bang Digital PLL With Transport Delay Using Fokker–Planck Equations," IEEE Trans. on Very Large Scale Integration (VLSI) Syst., 27(2), 398 –406, 2019. https://doi.org/10.1109/tvlsi.2018.2878665.

[11] W. Zhang, "An Analytical Solution of the Fokker-Planck Equation in the Phase-Locked Loop Transient Analysis," IEEE Trans. on Communications, 35(7), 773 – 775, 1987. https://doi.org/10.1109/tcom.1987.1096849.

[12] G. Hui, B. Chenghao, Z. Hongjun, L. Fanghui, L. Lei, "Synchronization Error Estimation of High-speed AD Sampling Based on Digital Phase Discrimination," in 2018 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC), Xuzhou, China, 2018. https://doi.org/10.1109/csqrwc.2018.8455461.

[13] H. Li, Z. Wang, H. Wang, "Research and implementation of high speed parallel carrier synchronization algorithm," in 2017 IEEE 9th International Conference on Communication Software and Networks (ICCSN), Guangzhou, China, 2017. https://doi.org/10.1109/iccsn.2017.8230073.

[14] F. Shaikh, B. Joseph, "Simulation of synchronous reference frame PLL for grid synchronization using Simulink," in 2017 International Conference on Advances in Computing, Communication and Control (ICAC3), Mumbai, India, 2017. https://doi.org/10.1109/icac3.2017.8318790.

[15] I. Park, I. Jang, W. Jung, S. Kim, "Noise immunity modeling and analysis of delay-locked loop," in 2015 IEEE 19th Workshop on Signal and Power Integrity (SPI), Berlin, Germany, 2015. https://doi.org/10.1109/sapiw.2015.7237401.

[16] C. Zhang, X. Zhao, X. Wang, X Chai, Zhang Z, X. Guo, "A Grid Synchronization PLL Method Based on Mixed Second- and Third-Order Generalized Integrator for DC Offset Elimination and Frequency Adaptability," IEEE Journal of Emerging and Selected Topics in Power Electronics, 6(3), 1517 – 1526, 2018. https://doi.org/10.1109/jestpe.2018.2810499.

[17] Q. Chen, M. Li, "Modified Gardner algorithm for bit synchronization in high-order QAM system," in 2013 International Conference on Computational Problem-Solving (ICCP), Jiuzhai, China, 2013. https://doi.org/10.1109/iccps.2013.6893575.

[18] T. Chang, K. Lee, G. Chen, S. Chiu, J. Yang, "Super resolution using trilateral filter regression interpolation," in 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP), Singapore, Singapore, 2017. https://doi.org/10.1109/siprocess.2017.8124511.

[19] J. Boiko, I. Kovtun, S. Petrashchuk, "Productivity of telecommunication systems with modified signal-code constructions," in 2017 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkov, Ukraine, 2017. https://doi.org/10.1109/infocommst.2017.8246374.

[20] W. Chen, M. Huang, X. Lou, "Sparse FIR Filter Design Based on Interpolation Technique," in 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP), Shanghai, China, 2018. https://doi.org/10.1109/icdsp.2018.8631685.

[21] S.K. Mitra, A. Mahalonobis, T. Saramaki, "A generalized structural subband decomposition of FIR filters and its application in efficient FIR filter design and implementation," IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing, 40(6), 363-374, 1993. https://doi.org/10.1109/82.277881.

[22] K. M. Whelan, F. Balado, N. J. Hurley, G. C. M. Silvestre, "A Two-Dimensional Extension of the Mueller and Müller Timing Error Detector," IEEE Signal Processing Letters, 14(7), 457-460, 2007. https://doi.org/10.1109/lsp.2006.891321.

[23] P. Harati, A. Dyskin, I. Kallfass, "Analog Carrier Recovery for Broadband Wireless Communication Links," in 2018 48th European Microwave Conference (EuMC), Madrid, Spain, 2018. https://doi.org/10.1109/eumc.2018.85.

[24] C. C. Cheng, S. Hikmet, S. Sezginer, Yu T. Su, "Enhanced Spatial Modulation With Multiple Signal Constellations," IEEE Trans. on Communications, 63(6), 2237 –2248, 2015. https://doi.org/10.1109/tcomm.2015.2422306.

[25] A. Wadhwa, U. Madhow, "Near-Coherent QPSK Performance With Coarse Phase Quantization: A Feedback-Based Architecture for Joint Phase/Frequency Synchronization and Demodulation," IEEE Trans. on Signal Processing, 64(17), 4432 –4443, 2016. https://doi.org/10.1109/tsp.2016.2568169.

[26] I.R. Parhomey, J.M. Boiko, O.I. Eromenko, "Features of digital signal processing in the information control systems of multipositional radar," Journal of Achievements in Materials and Manufacturing Engineering, 2(77), 75-84, 2016. https://doi.org/10.5604/17348412.1230101.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 6, 321-327 (2019)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

MIMO Fractional Order Control of a Water Tank Plant

Arturo Rojas–Moreno*, Juan Hernandez–Garagatti

Universidad de Ingenieria y Tecnologia - UTEC, Electronic Engineering Department, Lima 15063, Peru

ARTICLE INFO

Article history:
Received: 13 October, 2019
Accepted: 22 November, 2019
Online: 23 December, 2019

Keywords:
Fractional order controller
Integer order controller
Water tank plant
Centralized control

ABSTRACT

This work implements a MIMO (Multiple Input Multiple Output) FO (Fractional Order) control system for controlling the level and temperature of the water inside a tank by means of two inflow rates: cold and hot water, which are mixed to produce an outflow rate. Such a process exhibits coupling between inputs and outputs. A linear model of the plant is obtained experimentally. Such a model, the transfer matrix function of the plant, is used to design a centralized MIMO IO (Integer Order) controller that permits to achieve complete decoupling between the set points of level and temperature and the corresponding controlled outputs: level and temperature in the tank. The MIMO FO controller is obtained making fractional all de transfer functions of the MIMO IO controller. Experimental results demonstrate that the MIMO FO control system improves the control performance of the plant outputs: level and temperature of the water in the tank.

1 Introduction

In the literature, few studies have been performed for modelling a MIMO tank water plant having interaction (coupling) between its inputs: cold and hot water flow rates, and its outputs: level and temperature of the water into the tank. For instance, in [1], the level and temperature inside the tank are controlled by means of cold and hot water flow rates using two control configurations. The first configuration, called coupled control, employs two PID controllers, while the second, called decoupled control, uses two PID controllers and two decoupler devices. Figure 1 shows the controlled level and temperature using a coupled control configuration.

In [2], a water tank plant is modelled and controlled using two PID (Proportional Integral Derivative) controllers. The work in [3] employs a decoupled model of the water tank plant, which is controlled by means of two PID controllers and four decoupling devices, while in the work published in [4], the water tank plant is controlled by a MIMO PID controller using as control inputs the cold water flow rate and the electric current supplied to a heating resistance. The simulation work in [5] employs a fuzzy logic controller to control the temperature and water Level in a boiler. At the present, no work that employs a MIMO fractional order controller has been published.

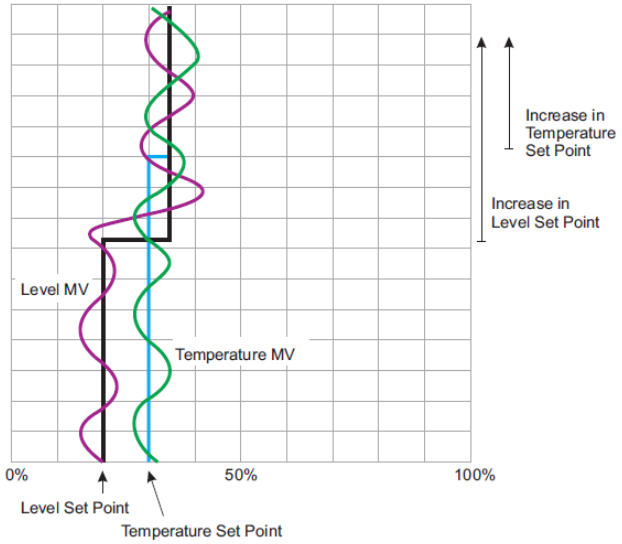


Figure 1: Controlled level and temperature using a decoupled control configuration. MV stands for Manipulated Variable. Taken from [1].

This work is organized as follows. Section 2 describes the multipurpose plant and the supervision module used in this work to implement the MIMO feedback control systems. Section 3 deals with the experimental modelling of the water tank plant. The design and implementation of a MIMO IO as well as a MIMO FO control systems are the topics of Sections

*Arturo Rojas–Moreno, Jr. Monte Algarrobo 518, Lima 15063, Peru, Phone (51) 948699637, Email: arojas@utec.edu.pe

4 and 5, respectively. Section 6 presents some concluding remarks derived from this work.

2 Plant Description

Figure 2 shows the multipurpose plant, patented by UTEC (Universidad de Ingenieria y Tecnologia) [6] used in this work, while Figure 3 depicts the corresponding supervision module patented by UTEC [7]. Such equipment is located in the Process Automation Lab of UTEC.



Figure 2: The multipurpose plant



Figure 3: The supervision module

Figure 4 depicts the P&ID (Piping and Instrument Diagram) of the multipurpose plant. In Figure 4, T–10 is the water tank plant. Observe that a flow rate q_C of cold water and a flow rate q_H of hot water are entering into T–10. A warm water flow rate q_D is exiting from T–10. Such a tank possesses a water overflow pipe not shown in Figure 4. The on–off valve DV–10 allows the passage of q_C. The control valve FV–10 is used to regulate q_C, while the flow transmitter FT–10 measures q_C.

The hot water flow rate of q_H is regulated by the control valve FV–11 and measured by the flow transmitter FT–11. Inside the tank T–10 there is a temperature transmitter TT–10 and a pressure transmitter PT–10, which is used as a level transmitter. In the tank T–10, there exists an electric heating resistance, whose electric current is controlled by the power controller PW–10 located in the supervision module. Both devices are not used in this paper. This work employs the tank T–20 to produce the required hot water flow rate q_H at a temperature of 50°C. The water flow rate q_D exits T–10 through the on–off valve DV–30.

Observe in Figure 4, that the on–off valve DV–20 permits the entering of the flow rate q_C to the tank T–20. This tank possesses a low level swith (LL–20) and a high level switch (HL–20) to indicate if the tank is either empty or filled with water, respectively. The water into the tank T–20 is heated electrically. The temperature into this tank is measured by the temperature transmitter TT–20 and controlled by means of the power controller PW–20 located in the supervision module. A flow rate q_H with a temperature of about 50°C is pumped to the tank T–10 by using the pump P–20. The pressure into the pipe that connects tanks T–10 and T–20 is measured by the pressure transmitter PT–20. The speed of the pump is kept constant by means of a speed controller (a variable frequency drive) located in the supervision module.

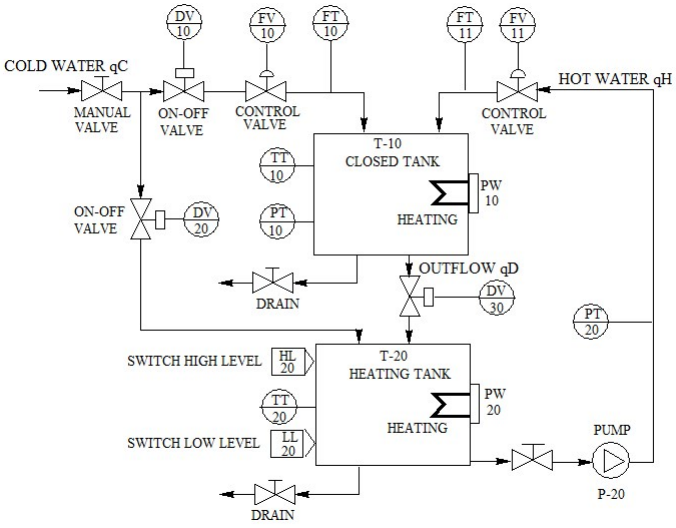


Figure 4: P&ID of of the multipurpose plant

Figure 3 shows the supervision module that is equipped with breakers, power sources, a PanelView front, two powerful PAC (Programmable Automation Controller), a PLC (Programmable Logic Controller), and a Flex I/O (Input/Output) switch among others. The latter device permits Ethernet communication between the PACs and the PLC of the supervision module with the valves and transmitters of the multipurpose plant. For such a purpose, input and output connectors available on the front panel of the supervisory module (Figure 3) permit to wire the PACs an PLC with the field instrumentation.

The software Studio 5000 from Rockwell Automation is used to elaborate the HMI (Human Machine Interface) and to

implement the control algorithms written in structured text language. This work employs the ControlLogic 5000 PAC. Table 1 shows the valued parameters and variables used in this paper.

Table 1: Parameters and variables employed in this work

Symbol	Description	Value
A	Tank rectangular section	0.12 m^2
A_o	Output pipe section	5.06×10^{-4} m^2
$\bar{h} = \bar{y}_1$	Steady state of h	0.2 m
$\bar{q}_C = \bar{u}_1$	Steady state of q_C	6.66×10^{-5} m^3/s
$\bar{q}_H = \bar{u}_2$	Steady state of q_H	7.66×10^{-5} m^3/s
$\bar{q}_D$	Steady state of q_D	12×10^{-5} m^3/s
g	Earth's gravity	9.81 m/s^2
θ_C	Temperature of q_C	289 K
θ_H	Temperature of q_H	321 K
$\bar{\theta} = \bar{y}_2$	Steady state of θ	304 K
ρ_C	Water density in q_C	998 kg/m^3
ρ_H	Water density in q_H	988 kg/m^3
ρ_D	Water density in q_D	995 kg/m^3
C_p	Heat specific capacity	4186.8 J/(kg–K)
C_d	Discharge coefficient	0.16
α	Discharge factor	3.586 $m^{2.5}/s$

The outflow rate q_D shown in Table 1 can be computed from

$$q_D = C_d\sqrt{2gh}D_d\frac{\pi}{4} = \alpha\sqrt{h} \quad \alpha = C_d\sqrt{2g}D_d \tag{1}$$

In (1), C_d is the dimensionless discharge coefficient, g is the gravitational acceleration, and D_d is the diameter of the pipe for the outflow q_D. Knowing the steady state values $\bar{h}$ and $\bar{q}_D$ of h and q_D, respectively,C_d and α can be calculated from

$$C_d = \frac{\bar{q}_D}{\sqrt{2g\bar{h}}D_d}; \qquad \alpha = C_d\sqrt{2g}D_d \tag{2}$$

Three experiments are performed to determine the discharge coefficient C_d shown in Table 1. For each experiment, the valve FV–10 is opened in a certain percentage to allow the flow rate q_C enter the tank T–10. At the same time, the flow rate q_D is regulated by means of the manual valve until the level h into the tank remains unchanged. At that point, the output flow rate of magnitude $\bar{q}_D$ equals the input flow rate of magnitude $\bar{q}_C$. Then, the discharge coefficient C_d can be computed from (2). Table 2 shows the results of the experiments. The selected value for C_d is 0.16.

Table 2: Experiment results to obtain C_d

$\bar{q}_C$ (m^3/s)	$\bar{h}$ (m)	C_d	α ($m^{2.5}/s$)
1.45× 10−4	0.162	0.160	3.586
1.95× 10−4	0.307	0.157	3.518
9.72× 10−5	0.092	0.142	3.182

3 Experimental Plant Modelling

Figure 5 depicts the block diagram of the MIMO LTI (Linear Time Invariant) control system, where s is the Laplace operator, $\mathbf{G}_p(s)$, $\mathbf{G}_c(s)$, $\mathbf{G}(s)$, and $\mathbf{G}_T(s)$ are transfer matrix functions of the plant, the controller, the open–loop system, and the closed–loop system, respectively. Also, **r**, **e**, **u**, and **y** are the reference, system error, control, and output vectors, respectively.

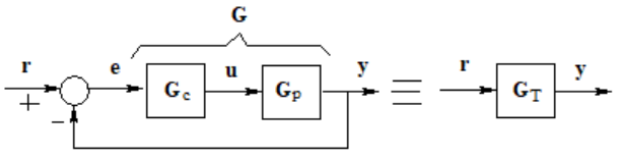


Figure 5: Block diagram of the MIMO feedback control system.

From Figure 5: $\mathbf{y}(s) = \mathbf{G}_p(s)\mathbf{u}(s)$, which can be expressed as

$$\begin{bmatrix} y_1(s) \\ y_2(s) \end{bmatrix} = \begin{bmatrix} G_{p11}(s) & G_{p12}(s) \\ G_{p21}(s) & G_{p22}(s) \end{bmatrix} \begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix} \tag{3}$$

The four transfer functions of (3) can be obtained experimentally from

$$\begin{aligned} G_{p11} &= \left[\frac{y_1}{u_1}\right]_{y_2=0,u_2=0} \quad G_{p12} = \left[\frac{y_1}{u_2}\right]_{y_2=0,u_1=0} \\ G_{p21} &= \left[\frac{y_2}{u_1}\right]_{y_1=0,u_2=0} \quad G_{p22} = \left[\frac{y_2}{u_2}\right]_{y_1=0,u_1=0} \end{aligned} \tag{4}$$

In the figures shown hereinafter, the water level and temperature in the tank are expressed in cm and ^{o}C, respectively. Let us consider a level transmitter span from 0 cm (0% of the span) to 40 cm (100% of the span). To determine G_{p11}, the valve FV–10 (Figure 4), which regulates the flow rate of cold water u_1, is opened from 30 to 50%, making y_1 (the water level into the T–10 tank) to change from 10 cm (25% of the span) to 19.5 cm (48.75% of the span) as seen in Figure 6. Using the tangent method in Figure 6, the transfer function G_{p11} with gain K_{p11} and time constant T_{p11} is found to be

$$G_{p11} = \left[\frac{y_1}{u_1}\right]_{y_2=0,u_2=0} = \frac{K_{p11}}{T_{p11}s+1} = \frac{1.1875}{180s+1} \tag{5}$$

$$K_{p11} = \frac{(48.75-25)}{(50-30)} = 1.1875$$

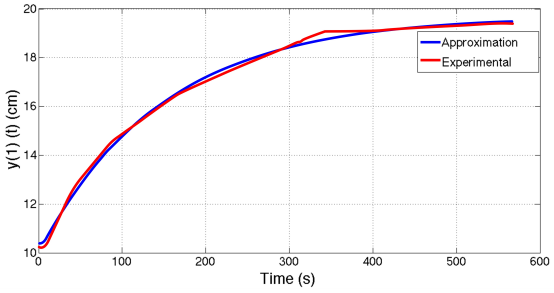


Figure 6: Experimental reaction curve for the G_{p11} transfer function.

To compute G_{p12}, the valve FV–11 (Figure 4), which regulates the flow rate of hot water u_2, is opened from 40 to 60%, making y_1 to vary from 10 cm (25% of the span) to 19.5 cm (48.75% of the span) as seen in Figure 7. Using the tangent method in Figure 7, the transfer function G_{p12} with gain K_{p12} and time constant T_{p12} is computed as

$$G_{p12} = \left[\frac{y_1}{u_2}\right]_{y_2=0,u_1=0} = \frac{K_{p12}}{T_{p12}s+1} = \frac{1.1875}{180s+1} \tag{6}$$

$$K_{p12} = \frac{(48.75-25)}{(60-40)} = 1.1875$$

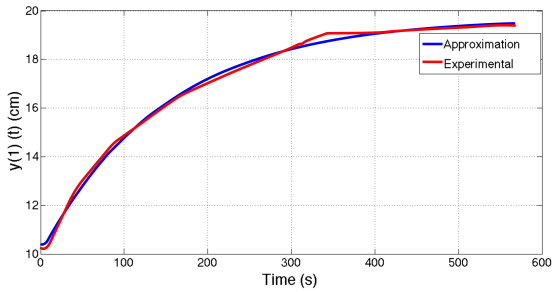


Figure 7: Experimental reaction curve for the G_{p12} transfer function.

Now, let us consider a temperature transmitter span from 16 ^{o}C (0% of the span) to 50 ^{o}C (100% of the span). To find G_{p21}, the valve FV–10, which regulates the flow rate of cold water u_1 is opened from 30 to 50%, making to drop the temperature y_2 from 32 ^{o}C (47% of the temperature span) to 24 ^{o}C (23.5%) as seen in Figure 8. From such a reaction curve, the transfer function G_{p21} with gain K_{p21} and time constant T_{p21} is calculated as

$$G_{p21} = \left[\frac{y_2}{u_1}\right]_{y_1=0,u_2=0} = \frac{-K_{p21}}{T_{p21}s+1} = \frac{-1.175}{180s+1} \tag{7}$$

$$K_{p21} = \frac{(47-23.5)}{(50-30)} = 1.175$$

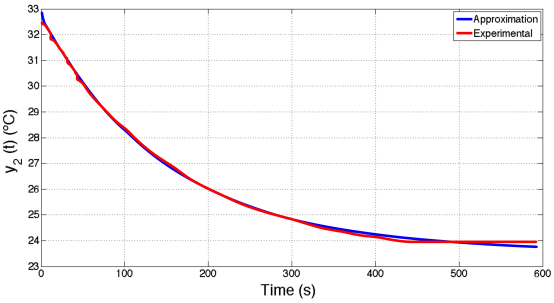


Figure 8: Experimental reaction curve for the G_{p21} transfer function.

To find G_{p22}, the valve FV–11 is opened from 40 to 60%, making to change y_2 from 33 ^{o}C (50% of the span) to 43 (79.4% of the span) as seen in Figure 9. From such a figure, the transfer function G_{p22} with gain K_{p22} and time constant T_{p22} is found to be

$$G_{p22} = \left[\frac{y_2}{u_2}\right]_{y_1=0,u_1=0} = \frac{K_{p22}}{T_{p22}s+1} = \frac{1.487}{180s+1} \tag{8}$$

$$K_{p22} = \frac{(79.74-50)}{(60-40)} = 1.487$$

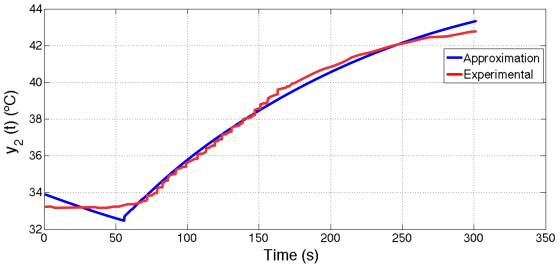


Figure 9: Experimental reaction curve for thr G_{p22} transfer function.

4 Design of the the Multivariable IO Control System

From Figure 5

$$\mathbf{G}(s) = \mathbf{G}_p(s)\mathbf{G}_c(s) \tag{9}$$

$$\mathbf{G}_T(s) = [\mathbf{G}(s) - \mathbf{I}]^{-1}\mathbf{G}(s) \tag{10}$$

Consider the following diagonal closed–loop transfer matrix function to assure complete decoupling between the different p inputs.

$$\mathbf{G}_T(s) = \begin{bmatrix} G_{T11} & & \\ & \ddots & \\ & & G_{Tpp} \end{bmatrix} \tag{11}$$

From (10)

$$\mathbf{G}(s) = \mathbf{G}_T(s)[\mathbf{I} - \mathbf{G}_T(s)]^{-1} \tag{12}$$

Since $\mathbf{G}$ is diagonal, $[I - \mathbf{G}]$ and $[I - \mathbf{G}]^{-1}$ are also diagonal matrices. Therefore, matrix $\mathbf{G}$ takes on the diagonal form

$$\mathbf{G}(s) = \begin{bmatrix} \frac{G_{T11}}{1+G_{T11}} & & \\ & \ddots & \\ & & \frac{G_{Tpp}}{1+G_{Tpp}} \end{bmatrix} \tag{13}$$

The system error is given by

$$\mathbf{e}(s) = \mathbf{r}(s) - \mathbf{y}(s) = [\mathbf{I} - \mathbf{G}_T(s)]\mathbf{r}(s) \tag{14}$$

The necessary condition to obtain $\mathbf{e}(t) = \mathbf{0}$ is

$$\lim_{s\to 0}\mathbf{G}_T(s) = \mathbf{I} \tag{15}$$

For instance, the following $\mathbf{G}_T(s)$ transfer matrix meets the condition given by (15)

$$\mathbf{G}_T(s) = \begin{bmatrix} \frac{1}{T_{11}s+1} & & \\ & \ddots & \\ & & \frac{1}{T_{pp}s+1} \end{bmatrix} \tag{16}$$

In (16), T_{ii} for $ii = 11, \cdots, pp$ are time constants. Introducing condition (15) into (10) results

$$\mathbf{I} + \mathbf{G}(0) = \mathbf{G}(0) \tag{17}$$

This requirement means that each element of the diagonal matrix $\mathbf{G}$ must contain at least one integrator. Using (9) into (10), we obtain the following MIMO controller

$$\begin{aligned} \mathbf{G}_c(s) &= [\mathbf{G}_p(s)]^{-1}\mathbf{G}_T(s)[\mathbf{I} - \mathbf{G}_T(s)]^{-1} \\ \mathbf{G}_c(s) &= [\mathbf{G}_p(s)]^{-1}\begin{bmatrix} \frac{1}{T_{11}s} & & \\ & \ddots & \\ & & \frac{1}{T_{pp}s} \end{bmatrix} \end{aligned} \tag{18}$$

Assuming that all parameters of $\mathbf{G}_p$ and $\mathbf{G}_T$ are known, then the following MIMO controller with known parameters takes on the form

$$\mathbf{G}_c(s) = \begin{bmatrix} K_{c11} + \frac{K_{i11}}{s} & -K_{c12} - \frac{K_{i12}}{s} \\ K_{c21} + \frac{K_{i21}}{s} & K_{c22} + \frac{K_{i22}}{s} \end{bmatrix} \tag{19}$$

Figure 10 depicts the simulation result of the MIMO IO control system using a sampling time of 0.1 s. Reference level and temperature signals were set to 20 cm and 33 ^{o}C.

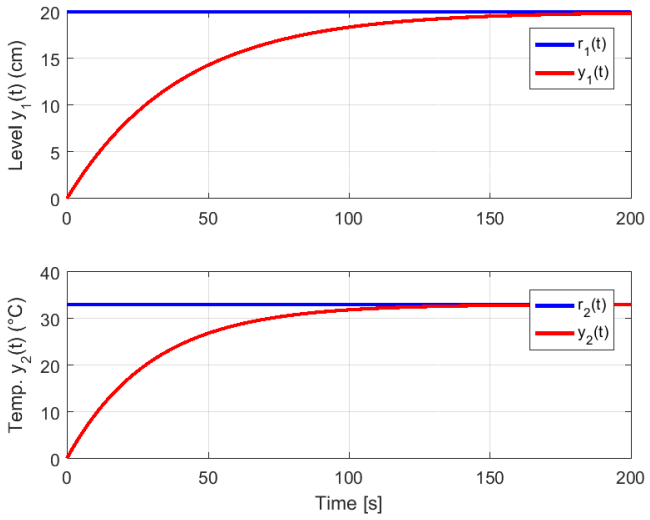


Figure 10: Simulated time–responses of the MIMO IO control system. Top graph: controlled level $y_1(t)$. Lower graph: controlled temperature $y_2(t)$.

Figure 11 depicts the experimental results of the MIMO IO control system employing most of the parameters obtained in the simulation phase. Some parameters required a real–time post–tuning to achieve the desired responses. Observe in Figure 11 that the controlled level $y_1(t)$ possesses a settling time of 225 s, null P.O. (Percent Overshoot), and about a null steady–state error. On the other hand, the controlled temperature $y_2(t)$ shows a settling time of 200 s, a P.O. of 25%, and around a null steady–state error.

It is worth to mention that the MIMO IO controller given by (19) constitutes the structure of the MIMO FO controller to be designed in the next Section.

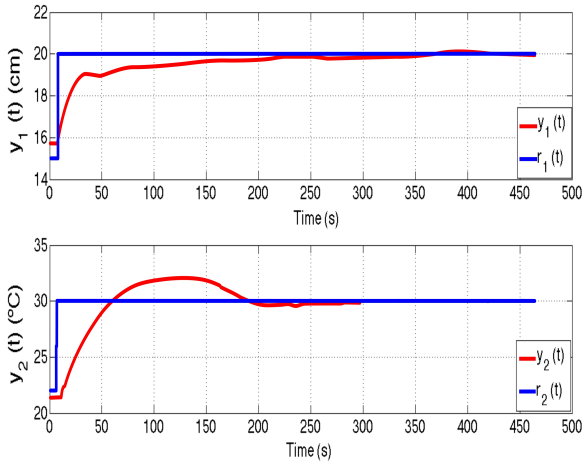


Figure 11: Experimental time–responses of the MIMO IO control system. Top graph: controlled level $y_1(t)$. Lower graph: controlled temperature $y_2(t)$.

5 Design of the MIMO FO Control System

The approach used in this section was employed in [9] to control two robot manipulators. The MIMO FO controllers is obtained making fractional the MIMO IO controller given by (19). That is, replacing in (19) all Laplace operators s by s^δ and s^λ, where δ and λ are fractional numbers between 0 and 1. Then, the MIMO FO controller takes on the form

$$\mathbf{G}_{cFO}(s) = \begin{bmatrix} K_{c11} + \frac{K_{i11}}{s^\delta} & -K_{c12} - \frac{K_{i12}}{s^\delta} \\ K_{c21} + \frac{K_{i21}}{s^\lambda} & K_{c22} + \frac{K_{i22}}{s^\lambda} \end{bmatrix} \tag{20}$$

The FO differentiators s^δ and s^λ given in (20) may be approximated by polynomials that depend on the Laplace operator s employing various formulas. For example, for the frequency range of operation $[\omega_b, \omega_h]$, s^δ and s^λ can be approximated by the following modified Oustaloup filter described in [10]

$$s^m \approx C \prod_{k=-N}^{N} \left(\frac{s + \omega_k^{'}}{s + \omega_k} \right) \tag{21}$$

$$C = \left(\frac{d\omega_h}{b} \right)^m \left[\frac{ds^2 + b\omega_h}{d(1-m)s^2 + b\omega_h s + dm} \right]$$

$$\omega_u = \sqrt{\omega_h/\omega_b}$$

$$\omega_k^{'} = \omega_b \omega_u^{(2k-1+m)/N} \qquad \omega_k = \omega_b \omega_u^{(2k-1-m)/N}$$

According to [10], the Oustalup filter produces a good approximation for $b = 10$ and $d = 9$. The frequency range of operation can be obtained from the Bode diagrams of transfer functions of the robot manipulator's transfer matrix function. However, such an approximation is not employed in this work because we will perform real–time implementation of recursive codes in the discrete–time domain.

There are various approximations for the FO differentiators s^{δ} and s^{λ} as a function of the shift operator z. This work employs the Muir's recursion method [11], which establishes

$$s^{\delta} \approx \left(\frac{2}{T}\right)^{\delta} \frac{A_n(z^{-1},\delta)}{A_n(z^{-1},-\delta)} \tag{22}$$

In (22), T is the sample time and z is the shift operator. $A_n(z^{-1},\delta)$ can be computed in recursive form as follows

$$\begin{aligned} A_n(z^{-1},\delta) &= A_{n-1}(z^{-1},\delta) - c_n z^{-n} A_{n-1}(z,\delta) \\ A_0(z^{-1},\delta) &= 1 \\ c_n &= \begin{cases} \delta/n & \text{if n is odd} \\ 1 & \text{0 if n is even} \end{cases} \end{aligned} \tag{23}$$

This work uses $n = 3$ in (23). Therefore

$$\begin{aligned} s^{\delta} &\approx \left(\frac{2}{T}\right)^{\delta} \frac{A_3(z^{-1},\delta)}{A_3(z^{-1},-\delta)} \\ A_3(z^{-1},\delta) &= -\frac{1}{3}\delta z^{-3} + \frac{1}{3}\delta^2 z^{-2} - \delta z^{-1} + 1 \\ A_3(z^{-1},-\delta) &= \frac{1}{3}\delta z^{-3} + \frac{1}{3}\delta^2 z^{-2} + \delta z^{-1} + 1 \end{aligned} \tag{24}$$

A similar expression to (24) is obtained for λ as follows

$$s^{\lambda} \approx \left(\frac{2}{T}\right)^{\lambda} \frac{A_3(z^{-1},\lambda)}{A_3(z^{-1},-\lambda)} \tag{25}$$

Using (20), we formulate $\mathbf{u} = \mathbf{G}_{cFO}\mathbf{e}$. Hence, the control laws u_1 and u_2 are formulated as

$$\begin{aligned} u_1 &= \left(K_{c11} + \frac{K_{i11}}{s^{\delta}}\right)e_1 - \left(K_{c12} + \frac{K_{i12}}{s^{\delta}}\right)e_2 \\ u_2 &= \left(K_{c21} + \frac{K_{i21}}{s^{\lambda}}\right)e_1 + \left(K_{c22} + \frac{K_{i22}}{s^{\lambda}}\right)e_1 \end{aligned} \tag{26}$$

In (26), $e_1 = r_1 - y_1$ and $e_2 = r_2 - y_2$ are the system errors, and r_1 and r_2 are the set points. Replacing (24) and (25) in (26), we obtain two control laws of the form

$$\begin{aligned} u_1(k) &= -\sum_{i=1}^{3} \alpha_i u_1(k-i) + \sum_{i=0}^{3} \beta_i e_1(k-i) + \\ &\quad \sum_{i=0}^{3} \rho_i e_2(k-i) \\ u_2(k) &= -\sum_{i=1}^{3} \eta_i u_1(k-i) + \sum_{i=0}^{3} \tau_i e_1(k-i) + \\ &\quad \sum_{i=0}^{3} \sigma_i e_2(k-i) \end{aligned} \tag{27}$$

In (27), k is the discrete time. Note that parameters α_i, β_i, and ρ_i depend on δ, while parameters η_i, τ_i, and σ_i depend on λ. Recall that fractional numbers δ and λ depend on the sampling time T.

Figure 12 depicts the simulation result of the MIMO FO control system obtained with the following parameters: T_{11} = 30, T_{22} = 40, K_{c11} = 1.012, K_{i11} = 0.005, K_{c12} = − 2.334, K_{i12} = − 0.013, K_{c21} = 0.0846, K_{i21} = 0.007, K_{c22} = 2.334, K_{i22} = 0.013. Reference level and temperature signals were set to 20 cm and 33 ^{o}C. The simulation phase employed a sampling time of 0.1 s.

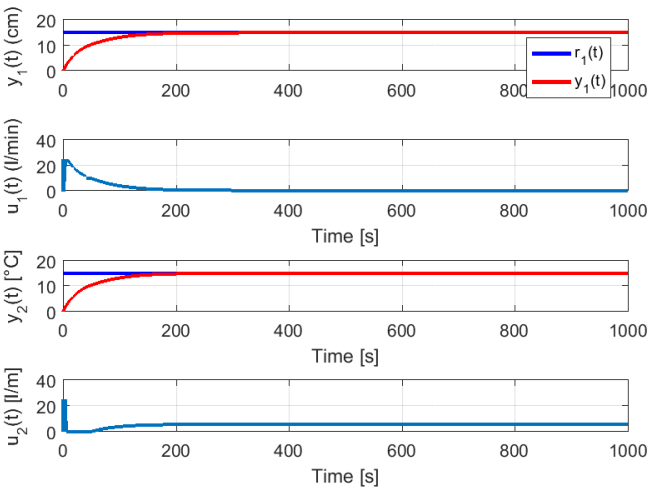


Figure 12: Simulation time–responses of the MIMO FO control system. Top graph: controlled level $y_1(t)$, second graph from the top: control force $u_1(t)$, third graph from the top: controlled temperature $y_2(t)$. Bottom graph: control force $u_2(t)$.

Figure 13 illustrates the experimental results of the MIMO FO control system using most of the parameters employed in the simulation phase. Some parameters needed a post-tuning to achieve the desired responses. Note in Figure 13 that the controlled level depicts a settling time of 70 s, null P.O., and null steady–state error. Also, the controlled temperature possesses a settling time of 125 s, null P.O., and null steady–state error.

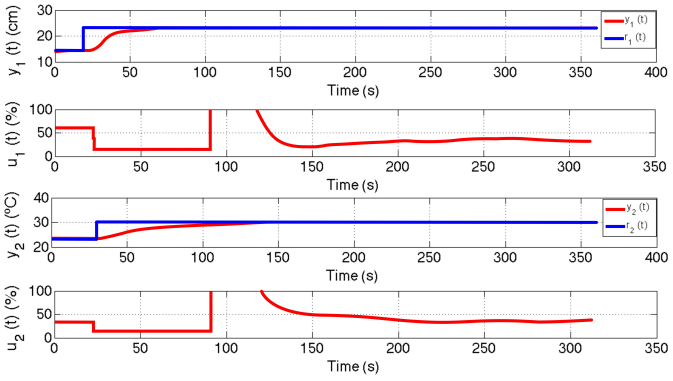


Figure 13: Experimental time–responses of the MIMO FO control system. Top graph: controlled level $y_1(t)$, second graph from the top: control force $u_1(t)$, third graph from the top: controlled temperature $y_2(t)$. Bottom graph: control force $u_2(t)$.

6 Concluding Remarks

A MIMO IO as well as a MIMO FO control systems were implemented for comparison purposes in this work. At the

present, no work that employs a MIMO fractional order controller has been published.

Experimental results demonstrate that the MIMO FO control system performs better because the settling time of the controlled level decreases from 225 s (Figure 11, top graph) to 70 s (Figure 13, top graph), while the settling time of the controlled temperature diminish from 200 s (Figure 11, lower graph) to 130 s (Figure 13, lower graph).

No P.O. (Percent Overshoot) shows the controlled level and temperature using a MIMO FO controller as seen in Figure 13. However, the controlled temperature using a MIMO IO controller depicts a P.O. of 20% as illustrated in the lower graph of Figure 11.

The main problem faced with the employed water tank plant was the supply of hot water in sufficient quantity to perform the experiments.

The simulation of the MIMO IO control systems is necessary to analyse the behaviour of the controlled plant and estimate the tuning parameters required for real–time implementation.

As seen in Figures 11 and 13, the controlled level and temperature do not present oscillations. However, using the decoupled control configuration developed in [1], the controlled level and temperature depicted in Figure 1 show strong oscillations.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment We are very grateful to the Electronic Engineering Department of the Universidad de Ingenieria y Tecnologia (UTEC) for supporting this research work.

References

[1] TE37 Equipement Control and Instrumentation Study Station, User Guide, Tecquipment Ltd, UK, 2009.

[2] V. Tzouanas, "Temperature and Level Control of a Multivariable Water Tank Process", in 120th ASEE Annual Conference & Exposition, Atlanta, Georgia, Jun 23–26, 2013, pp. 1–11.

[3] E. Cornieles, et al., "Modelling and Simulation of a Multivariable Process Control", in IEEE ISIE, Montreal, Quebec, Canada, July 9–12, 2006, pp. 1–11. DOI: 10.1109/ISIE.2006.296039

[4] A Rojas–Moreno and A. Parra–Quispe, "Design and Implementation of a Water Tank Control System using a Multivariable PID Controller", in 6th International Conference on Computing, Communications and Control Technologies (CCCT), Orlando, USA, June 2008. ISBN: 1934272442 9781934272442, Number OCLC: 552110236

[5] C. Dinakaran, "Temperature and Water Level Control in Boiler by using Fuzzy Logic Controller", International Journal of Electrical Power System and Technology, Vol. 1: Issue 1, pp. 27–37, 2015

[6] A. Rojas–Moreno, et al.,"Planta industrial multipropósito para control e instrumentación." Resolution N° 001545–2019/DIN-INDECOPI, May 20, 2019, Lima, Peru

[7] A. Rojas–Moreno, et al., "Planta industrial multipropósito para control y supervisión." Resolution N° 001035–2019/DIN-INDECOPI, March 28, 2019, Lima, Peru.

[8] A. Rojas–Moreno and J. Hernandez–Garagatti, "Modelling of a Multipurpose Water Tank Plant", in IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON), Cusco, Peru, Aug. 15–18, 2017, pp. 1–4. https://doi.org/10.1109/intercon.2017.8079668

[9] A. Rojas–Moreno, "An Approach to Design MIMO FO Controllers for Unstable Nonlinear Plants," IEEE/CAA Journal of Automatica Sinica, vol. 3, no. 3, July 2016, pp. 338–344. https://doi.org/10.1109/jas.2016.7508810

[10] C. A. Monje et al., "Implementation of Fractional Order Controllers," in Fractional–Order Systems and Control. Fundamentals and Applications, Springer–Verlag London Limited 2010, pp. 195–196.

[11] B. Vinagre, et al., "Two direct Tustin discretization methods for fractional-order differentiator/integrator," J. of the Franklin Institute, 340, pp. 349–362, 2003. https://doi.org/10.1016/j.jfranklin.2003.08.001

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 2, 324-333 (2020)
www.astesj.com

ASTES Journal
ISSN: 2415-6698

ANN Based MRAC-PID Controller Implementation for a Furuta Pendulum System Stabilization

Efrain Mendez*, German Baltazar-Reyes, Israel Macias, Adriana Vargas-Martinez, Jorge de Jesus Lozoya-Santos, Ricardo Ramirez-Mendoza, Ruben Morales-Menendez and Arturo Molina

Tecnologico de Monterrey, School of Engineering and Sciences, Mexico

ARTICLE INFO

Article history:
Received: 15 January, 2020
Accepted: 09 May, 2020
Online: 30 May, 2020

Keywords:
Furuta Pendulum
Robotics
Robotic systems
UMS
Artificial Intelligence
ANN
MRAC-PID
Non-linear Systems

ABSTRACT

Nowadays, process automation and smart systems have gained increasing importance in a wide variety of sectors, and robotics have a fundamental role in it. Therefore, it has attracted greater research interests; among them, Underactuated Mechanical Systems (UMS) have been the subject of many studies, due to their application capabilities in different disciplines. Nevertheless, control of UMS is remarkably more difficult compared to other mechanical systems, owing to their non-linearities caused by the presence of fewer independent control actuators with respect to the degrees of freedom of the mechanism (which characterizes the UMS). Among them, the Furuta Pendulum has been frequently listed as an ideal showcase for different controller models, controlled often through non-lineal controllers like Sliding-Mode and Model Reference Adaptive controllers (SMC and MRAC respectively). In the case of SMC the chattering is the price to be paid, meanwhile issues regarding the coupling between control and the adaptation loops are the main drawbacks for MRAC approaches; coupled with the obvious complexity of implementation of both controllers. Hence, recovering the best features of the MRAC, an Artificial Neural Network (ANN) is implemented in this work, in order to take advantage of their classification capabilities for non-linear systems, their low computational cost and therefore, their suitability for simple implementations. The proposal in this work, shows an improved behavior for the stabilization of the system in the upright position, compared to a typical MRAC-PID structure, managing to keep the pendulum in the desired position with reduced oscillations. This work, is oriented to the real implementation of the embedded controller system for the Furuta pendulum, through a Microcontroller Unit (MCU). Results in this work, shows an average 58.39% *improvement regarding the error through time and the effort from the controller.*

1 Introduction

In recent decades, robotics have gained increasing importance in endless applications for different disciplines (as explained in [1]); such as, robotic manipulators for industrial automation, precision robots to perform surgeries, automation of assembly lines, among others. Thus, robotics have attracted greater interest multidisciplinary researchers.

Moreover, according to [2], it is defined in robotics that an Underactuated Mechanical System (UMS), is a scheme with more Degrees of Freedom (DOF) compared to its control actuators. Additionally, [3] explains that the control of UMS's is a very active research topic, due to their broad applications in Robotics, Aerospace, and Marine Vehicles.

Therefore, in spite of the complexity caused by the lack of actuators to control the movement of the system, it is precisely the low number of required actuators that makes UMS's ideal for applications where energy efficiency is sought. Additionally, UMS's are systems that allow to decrease the size of the manipulators, and even simplify the amount of elements of a more complex system. The above, resulting also in cost reduction with an increased process efficiency, which is the key element that attracts the most interest for its development in the industry.

The non-linearities caused by the relation between the actuators and the DOF to be controlled, makes the complexity of these systems attractive as testbeds, for the research of different control structures. Among them, this work is focused in the rotatory pendulum, also known as the *Furuta Pendulum* (in honor of its inventor *K.*

*Corresponding Author: Efrain Mendez Flores; Tecnologico de Monterrey, School of Engineering and Sciences, Mexico City 14380, Mexico; Correspondence: efrain@tec.mx; Tel.: +52-5521-333-361.

Furuta [4]), which is a highly non-linear system that is unstable at the desired upright position, as described in [3].

Hence, different solutions have been proposed addressing the swinging-up and the upright issues of the Furuta Pendulum, which are the main control objectives to fulfill in order to achieve the desired position. In the case of the stabilization issue, [5] proposes a Linear Quadratic Regulator (LQR) approach, showing a better settling time with low overshoots, regarding a Sliding-Mode Controller (SMC). Nevertheless, solutions that require linearized equations are more effective when the non-linear components of the system, are small regarding the linear predominant behavior; therefore, in high non-linear systems like the Furuta Pendulum, the control structure may be highly sensitive to parametric variations, compromising the behavior of the controller when additional friction coefficients, unexpected center of mass changes in the real, or even the uncertainty of the linearized model, are presented.

Therefore, many other approaches based on nonlinear controllers are still being studied, as shown in [6] and [7] applications, where the implementation of the SMC and an Integral Sliding Mode Controller (ISMC) for pendulum stabilization are presented respectively, showing that the (ISMC) has the best response regarding perturbations rejection. Nevertheless, the chattering continues to be a limitation regarding to practical limitations of actuators, as explained in [8].

On the other hand, [9] presents a Model Reference Adaptive Controller (MRAC) implementation, discussing the implementation of the technique compared again to an LQR approach. In this particular case, it is shown that the LQR controller had an improved behavior regarding the settling time of the MRAC controller through the MIT rule, nevertheless perturbation rejection tests where not clearly analyzed.

Additionally, other studies such as [10], validates that the adaptive features of the MRAC can create a more robust system, features that are also translated to all the MRAC derivations. The previous premise can be clearly validated by the work presented in [11], where an MRAC combined with a Proportional Integrative Derivative (MRAC-PID) is implemented in a real Furuta pendulum testbed, achieving through the simulated and the experimental responses the validation of the method, for the upright stabilization issue of the pendulum.

Henceforth, it is important to highlight that this paper is an extension of the [11] work, originally presented in the *ICMEAE 2019* (6th International Conference on Mechatronics, Electronics and Automatation Engineering) and *ICCRE 2019* (4th International Conference on Control and Robotics Engineering); where the design, modelling and a first control approach for the Furuta Pendulum was presented.

Nevertheless, the [11] implementation had its main drawback, in the great amount of processing resources consumed from the Microcontroller Unit (MCU), which in some instances compromised the correct and precise sampling rate of the data, which leaded to unexpected noise in the error signal and therefore to practical issues in the implementation. Therefore, this proposal takes the obtained results from the MRAC-PID approach, improving the behavior of the system through the implementation of an ANN based controller.

Controllers based on ANN are able to provide robust, non-fluctuating performances in the presence of parametric uncertainty and extraneous disturbances, as explained in [12] where an ANN was simulated in order to validate its capabilities for UMS. However, the methodology was only simulated, and therefore there was none experimental validation of the performance of the control theory.

On the other hand, the main objective of designed ANN in this work, is to emulate the behavior of the MRAC-PID controller implemented in [11], extending the operating margins and maintaining the stability gained by the original MRAC-PID, but with lower computational effort; which is indeed, a common characteristic of the implementation of ANN's (as discussed in [13] and [10]).

In this particular proposal, the ANN is not trained through a model oriented approach as in [12]; in this work, the model of the Furuta pendulum was taken for the MRAC-PID controller calibration and implementation, and the ANN was trained through the acquired data from the experimental testbed, expecting to achieve a correct emulation of the controller's behavior, but improving its performance by increasing response speed of the controller through the reduction of the required embedded processing effort.

Then, the results in this work where acquired through the embedded implementation in the real experimental testbed. It is critical for this work to highlight that, the ANN is trained through the input/output data acquired from the stabilization MRAC-PID controller implemented in [11], in order to achieve the improved behavior of the controller with less computational effort.

Therefore, the control effort is validated through the evaluation of the control signal through time, where the duty cycle sent to the driver of the actuator can be evaluated for the task. Meanwhile, the performance is evaluated by the evolution of the error through time. Together, both signals allow evaluating the performance of the system and, the actuator stress required to achieve the control objective.

Hence, this work presents in Section 2 the control objectives of the Furuta Pendulum with greater details, meanwhile the experimental testbed is presented in Section 3; Section 4 discusses the structure of the MRAC-PID control topology, leading to the ANN concepts discussed in Section 5. Finally, Section 6 presents the results of this work and Section 7 presents the final conclusion.

2 Furuta Pendulum

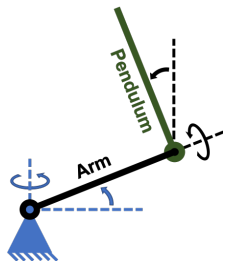


Figure 1: Basic Furuta Pendulum diagram.

As explained in [14], the rotatory inverted pendulum (also known as the Furuta Pendulum), is a rotatory system composed by an homogeneous pendulum attached to a perpendicular arm, which rotates in a horizontal plane. Fig. 1 shows the basic configuration of the pendulum, where the blue triangle represents the fixed point in which

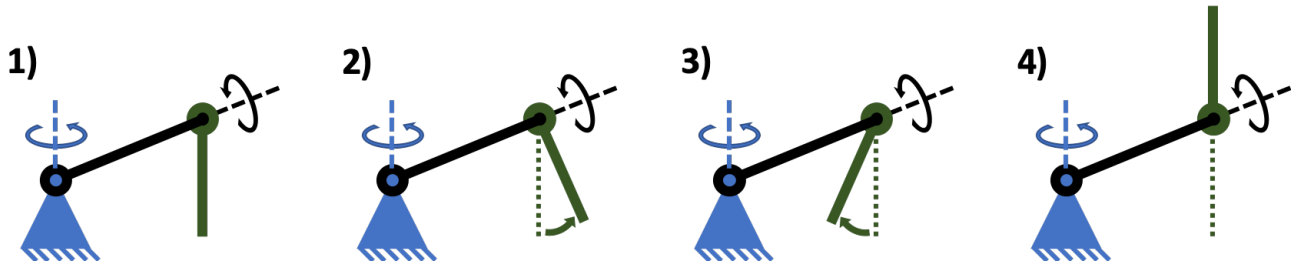


Figure 2: Swinging-up sequence for the Pendulum.

the arm is attached, therefore the first DOF of the system is in the arm, that can only rotate around the vertical axis, in the horizontal plane.

Moreover, also from Fig. 1 can be seen that the pendulum is coupled to the arm by means of a bearing, which allows it to rotate around the arm axis and therefore, it is there where the next degree of freedom of the system is found. Henceforth, attending to the UMS classification in which the Furuta Pendulum is located in, the only actuator of the system is attached in the fixed point at the beginning of the arm.

2.1 Control Objectives

As implied from the structure of the Furuta Pendulum, it is clear that it is a system that is unstable without a controller (as also explored in [15]). Thus, the control objectives of a Furuta Pendulum can be divided into two main issues, as addressed in [11], where they have been extensively studied both individually and together.

On the one hand, the initial state of the Furuta pendulum is just as shown in the first diagram of Fig. 2, where the pendulum is in its resting position. The control objective there, is located in the natural equilibrium point; in this scenario, a controller needs to be applied in order to destabilize the system by applying torque to the arm through the actuator, traduced into a swing-up routine, by then, arising the pendulum into the working range of the main stabilization controller.

The swing-up routine can be summarized by the second and third diagrams of Fig. 2, finally leading to the stage of the fourth diagram, where the pendulum switches to the controller in charge of maintaining the upright position. The swinging-up routine can be made through energy control, as explored by Furuta itself in [4]. Other proposed solutions for the task is the predictive controller ([16]), and even solved through feedforward approaches such as presented in [17].

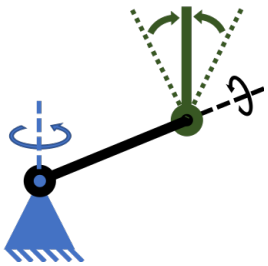


Figure 3: Representation of the stabilization objective.

On the other hand, the second control objective lies in the stabilization of the pendulum, in order to maintain it in an upright position by correcting any disturbance, only through the applied torque of the arm. Fig. 3, shows the basic outline of the control objective, where the green doted lines represent the operating range that delimits the control actions, knowing that most of the state of the art recommends it to be maximum $\pm 15^o$ from the upright position.

Therefore, summarizing the control objectives, out of the stabilization range the swing-up controller is activated, in order to take the pendulum into the area in which the stabilization controller takes the wheel, achieving then the expected upright position.

In addition, the study of controllers in charge of maintaining the pendulum in the upright position, have served for inspiration of many applications, such as the stability of walking humanoid robots among others. Hence, this work proposal is focused in a solution regarding the stabilization control objective, where ANN based controller seeks to achieve an effective solution for its microcontroller implementation.

2.2 System Modelling

As explained in [11], the mathematical model of the system can be delivered by the segmentation of the mechanical system, through its kinetic and potential energy (as discussed in [4]. Therefore, the energies relation from which the mathematical model of the pendulum emerges, is given by (1).

$$L = \{K_{arm} + K_p\} + \{V_{arm} + V_p\} \tag{1}$$

where L is the Lagrange relation obtained by combining both energy equations, K_{arm} and k_p are the kinetic energy equations of the arm and the pendulum respectively, meanwhile V_{arm} and V_p are their potential energy equations. After the kinetic and potential energies analysis, as explained in [18] and also validated by [11], the expansion of (1) leads to second order differential equations; the first one, is the differential equation of the motion of the pendulum in terms of the angle generated between the horizontal plane and the pendulum's arm (θ_{arm}). Meanwhile, the second one is the differential equation obtained from the motion of the system, in terms of the pendulum and the angle between the vertical axis and the pendulum itself (θ_p).

The simplified expression that describes the behavior of (θ_{arm}), is given by (2). Where α and β are products of a change of variable, since α is the expansion given by (3), and β is developed through (4).

$$\ddot{\theta}_{arm} = \frac{\alpha}{\beta} \tag{2}$$

$$\begin{aligned}\alpha = \tau\,(m_p\,{l_p}^2 + I_p) - \sin(\theta_p)\,({l_p}^3\,l_r\,{m_p}^2\,\dot{\theta}_p^2\cos(\theta_p)^2\\ -\dot{\theta_p}^2\,(l_r\,{l_p}^3\,{m_p}^2 + I_p\,l_r\,l_p\,m_p) + g\,{l_p}^2\,l_r\,{m_p}^2\,\cos(\theta_p))+\\ \dot{\theta}_p\,\sin(2\,\theta_p)\,(\dot{\theta}_{arm}\,{l_p}^4\,{m_p}^2 + l_p\,\dot{\theta}_{arm}\,{l_p}^2\,m_p)\end{aligned} \tag{3}$$

$$\begin{aligned}\beta = m_p\,(I_r\,{l_p}^2 + l_p\,{l_r}^2) + \sin(\theta_p)^2\,({l_p}^4\,{m_p}^2 + l_p\,{l_p}^2\,m_p)\\ +l_p\,I_r + {l_p}^2\,{l_r}^2\,{m_p}^2 - {l_p}^2\,{l_r}^2\,{m_p}^2\cos(\theta_p)^2\end{aligned} \tag{4}$$

On the other hand, the behavior of (θ_p) is given by (5). Where another change of variable was made in order to simplify the expression, as explored in (2). Therefore in this case, γ is given by the expression given by (6), meanwhile λ is traduced as (7).

$$\ddot{\theta_p} = \frac{\gamma}{\lambda} \tag{5}$$

$$\begin{aligned}\gamma = \sin(\theta_p)^3\,(\cos(\theta_p)\,{l_p}^4\,{m_p}^2\,\dot{\theta}_{arm}^2 + g\,{l_p}^3\,{m_p}^2)+\\ \sin(\theta_p)\,(\cos(\theta_p)\,({l_p}^2\,{lr}^2\,{m_p}^2\,\dot{\theta}_{arm}^2\\ -{l_p}^2\,{lr}^2\,{m_p}^2\,\dot{\theta_p}^2 + I_r\,{l_p}^2\,m_p\,\dot{\theta}_{arm}^2)\\ +g\,l_p\,{lr}^2\,{m_p}^2 + I_r\,g\,l_p\,m_p) - l_p\,lr\,m_p\,\tau\,\cos(\theta_p)\\ +{l_p}^3\,lr\,{m_p}^2\,\dot{\theta}_{arm}\,\dot{\theta}_p\,\sin(2\,\theta_p)\,\cos(\theta_p)\end{aligned} \tag{6}$$

$$\begin{aligned}\lambda = m_p\,(I_r\,{l_p}^2 + I_p\,{lr}^2) + \sin(\theta_p)^2\,({l_p}^4\,{m_p}^2 + I_p\,{l_p}^2\,m_p)\\ +I_p\,I_r + {l_p}^2\,{lr}^2\,{m_p}^2 - {l_p}^2\,{lr}^2\,{m_p}^2\cos(\theta_p)^2\end{aligned} \tag{7}$$

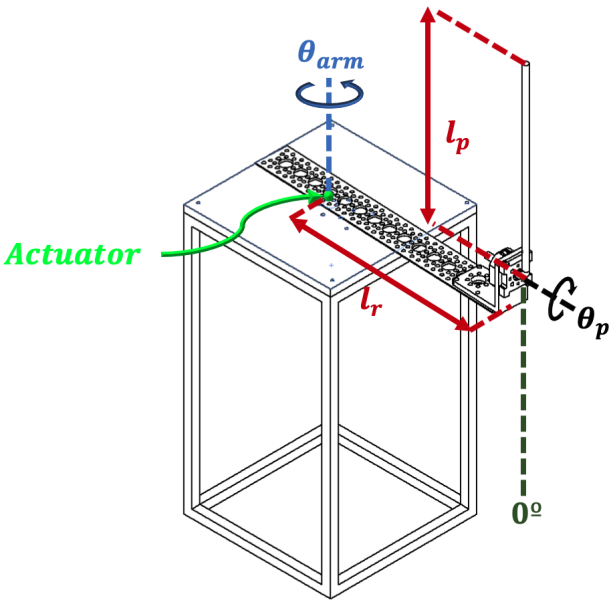


Figure 4: Representation of the stabilization objective.

In terms of the testbed analyzed in this paper, the variables from (2) and (5) are sketched in the diagram shown in Fig. 4, where an isometric view of the designed 3D model of the pendulum is used, design that will be explained in more detail later in this document.

Nevertheless, all the variables of the equations presented in this section and analyzed in Fig. 4, are summarized in Table 1. Additionally, Table 1 shows the parameters of the each variable in the implemented experimental testbed, which is explained in Section 3.

Table 1: Furuta Pendulum Variables

Variable	*Description*	*Value*
m_r	Arm mass	0.35 [kg]
m_p	Pendulum mass	0.12 [kg]
l_r	Arm length	0.15 [m]
l_p	Pendulum length	0.25 [m]
I_r	Arm moment of inertia	0.02 [kgm^2]
I_p	Pendulum moment of inertia	0.001 [kgm^2]
τ	Nominal Torque	0.09 [Nm]
θ_{arm}	Arm angle	- [rad]
θ_p	Pendulum angle	- [rad]

It is important to highlight that in this application, the arm length l_r is measured as the distance from the center of the actuator shaft to the location of the pendulum. On the other hand, the length of the pendulum is taken as the length of the pendulum bar, from the point where it is attached to the system. Nevertheless, the center of mass of the pendulum is found as the half of the pendulum's length.

3 Experimental Testbed

The designed testbed, was inspired by the one implemented in the short version of this work ([11]), where basically the mechanical design is described in Fig. 5.

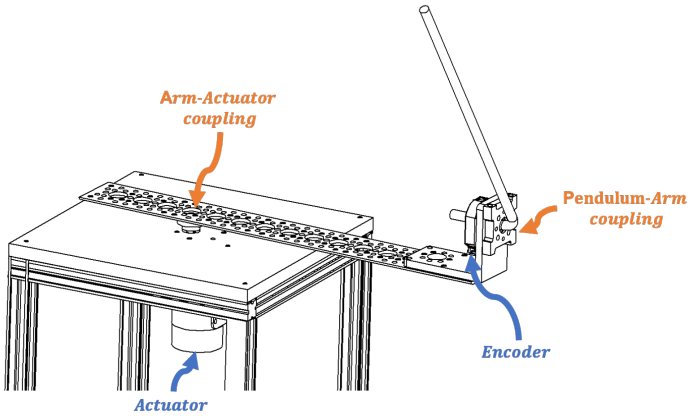


Figure 5: Schematic of the component assembly for the testbed.

The selected actuator, was the *Pololu 12V, 19:1 Gear Motor w/ 64 CPR Encoder*, which is a DC motor with a gearbox, that already has a two phase encoder for speed measurements that also has a 5*A* stall Current, same that is activated and controlled thorough a *MD30C* 30*A* DC Motor Driver.

Additionally, the control signal of the system, the signals processing and data acquisition, are made by the *FRDM-K64F* development board, which has an *MK64FN1M0VLL12* microcontroller that

works at 120MHz, with 1024KB of Flash memory, 256KB SRAM memory, and also allows an easier implementation of the equations through its integrated Floating Point Unit (as explored in [11]).

It is important to highlight that the original version of this work ([11]), was programmed using *FreeRTOS* (Free Real-Time Operating System), in order to work with a better sample time response from the microcontroller and a faster control output signal. Nevertheless, several changes were made in order to improve the sampling rate and therefore the error post-processing. The main changes, lie in the implementation of a Periodic Interrupt Timer (PIT), which generates highly precise interrupts at regular intervals with minimal processor intervention, therefore allowing to exploit more resources for the controller implementation.

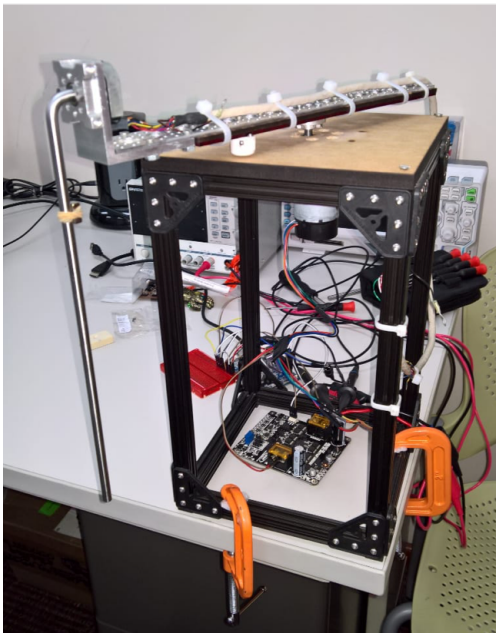


Figure 6: Experimental testbed designed and implemented.

On the other hand, for the pendulum's data acquisition, a modular-incremental encoder was used to acquire the data using the PIT, in order to determine the speed and the position signals of the pendulum. The encoder *Cui-AMT113Q* (also highlighted in Fig. 5) was selected for the task, since it is a quadrature encoder with two phases configured with 1024 counts per revolution.

Therefore with the selected hardware, the measurement of θ_p has a resolution of 0.35^o, which allows monitoring the pendulum's position with a correct error estimation for the controller implementation.

Finally, Fig. 6 shows the experimental testbed implemented for the controller evaluation, where the implemented hardware was taken as shown in Fig 5. It can be noted in Fig. 6 in black, the frame that structurally supports the system is observed, with a wooden surface where the motor is attached. Consequently, as mentioned above the arm is directly coupled to the motor shaft, and finally in the arm the bearing that holds the pendulum is attached together with the encoder.

4 MRAC-PID controller

As highlighted all along this work, the main objective addressed in this work is to maintain a steady-state around the upright position of the pendulum. Therefore, as analyzed and validated in [11], the position of the pendulum can be controlled through a Model Reference Adaptive Controller (MRAC) with a Proportional Integrative Derivative (PID) Controller. Thus, the structure of the controller implemented in [11], was inspired by the topology shown in [19].

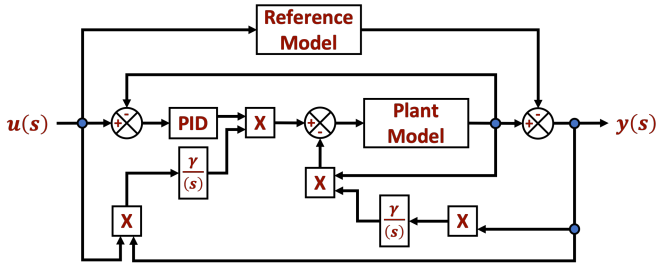


Figure 7: Blocks diagram of the MRAC-PID structure.

Then, Fig. 7 shows the general structure of the implemented controller, where the PID block from the diagram is taken as the classic PID parallel topology, which is represented by Fig. 8.

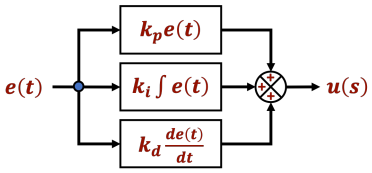


Figure 8: Blocks diagram of the Parallel PID structure.

Henceforth, the PID structure of the controller described by Fig. 8, can be expressed as shown by (8).

$$U(t) = k_p e(t) + k_i \int e(t)dt + k_d \frac{de(t)}{dt} \tag{8}$$

where k_p, k_i and k_d are the proportional, integral and derivative gains respectively, and $e(t)$ is the error signal. On the other hand, the reference model described in the blocks diagram from Fig. 7, is taken as a general second order transfer function, which is implemented through the expression defined in (9).

$$ReferenceModel = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{9}$$

where, ζ is the dampening ratio of the reference system, which can be delivered through (10).

$$\zeta = \frac{-\ln(M_p)}{\sqrt{\pi^2 + \ln(M_p)}} \tag{10}$$

where in this case, M_p is taken as the maximum allowed overshoot of the system. Additionally from (9), ω_n is the natural frequency of the reference model, which can be defined by (11).

$$w_n = \frac{4.6}{\zeta * t_s} \tag{11}$$

where t_s represents the settling time. Moreover, for this application, the settling time is taken as $t_s = 0.001$, with a maximum overshoot $M_p = 0.02$, where the learning rate for the MRAC is $\gamma = 0.25$.

Nevertheless, according to the validation shown in [11], the proposed MRAC-PID controller can be implemented in embedded systems, by using the equations developed in [19], where the constants allow real-time tuning of the controller that learns from the actual behavior of the error. Therefore, the initially implemented equations in the embedded system are:

$$k_p = \frac{-\gamma * err * (u_c - y) * w_n}{p^2 + 2\zeta w_n p + w_n^2} \tag{12}$$

$$k_i = \frac{-\gamma * err * (u_c - y) * w_n^2}{(p^2 + 2\zeta w_n p + w_n^2) * p} \tag{13}$$

$$k_d = \frac{-\gamma * err * y * p}{p^2 + 2\zeta w_n p + w_n^2} \tag{14}$$

where, err is the normalized error of the system, u_c the output of the system, y the reference model output, and p is the slope defined by $p = \frac{d(err)}{d(t)}$.

Consequently, in order to validate the behavior of the controller from which the ANN is going to learn from, Fig. 9 shows the the pendulum using the proposed control topology, which as validated before in [11], it is clear the capabilities of the MRAC-PID to maintain the pendulum in a steady-state in the upright position.

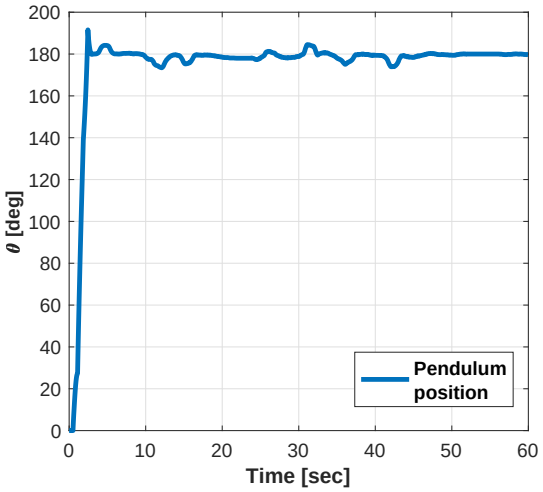


Figure 9: Validation through pendulum's position against time.

Basically, the Fig. 9 was delivered after sampling a minute of the pendulum's behavior, using a 50[$mSec$] sampling rate. In every performed test for this work, the pendulum starts at a 0° initial point (as shown in Fig. 4), and the reference is in the upright position at 180°. Subsequently, with the control objective already fulfilled, the next stage of this work is to take a step into the design based on ANN.

5 Artificial Neural Networks

Artificial Neural Networks (ANN) were first introduced by McCulloch and Pitts in [20] as a propositional logic model that intended to imitate biological neurons. The main objective of their proposal was to use models of artificial neurons, in order to evaluate any logical computation based on AND, OR, and NOT assignments.

Nevertheless, the interconnection of multiple neuron models as seen in Fig. 10, enables to assign different weights at each connection, which allows a better prediction of the training model through the implementation of backpropagation models between them.

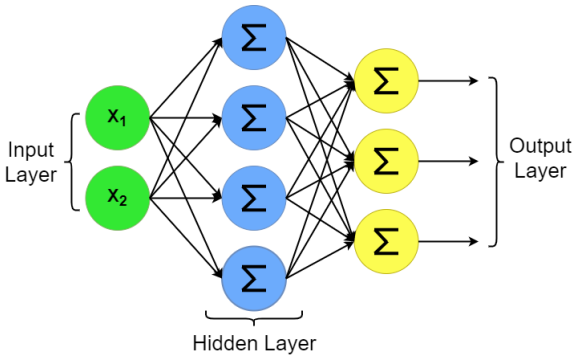


Figure 10: Multiple-layer neuron model

Therefore, knowing that it is possible to use multi-layer neural networks for regression purposes, it can be designed an ANN according to the input-output relationship of the control system (as explained in [21]). Moreover, in the case of a Multiple Input-Single Output (MISO) model, the input layer of the network will assign one neuron to each input, while the final layer will include only one single neuron for the output.

Therefore, regarding the layers of the network, it is commonly used a Rectified Linear Unit (ReLU) function for the hidden layers, which can be described by the expression shown in (15).

$$y = max(0, x) \tag{15}$$

Meanwhile, any linear equation can be used to guarantee positive values; although the selection of the function could vary depending on the nature of the system, for this implementation a Sigmoid given by (16)), was selection for the activation task.

$$y = \frac{1}{1 + e^{-x}} \tag{16}$$

For the regression, the network loss function seeks to minimize the error from the ANN, and can be performed (as explored in [22, 23, 24]) through the mean-squared error (MSE) as given by (17).

$$L = E\left[\|\epsilon\|^2\right] = E\left[\|y - f(x)\|^2\right] \tag{17}$$

where the estimation of $f(x)$ can be expressed as the product of the input values with the networks weights, as seen in (18).

$$\hat{y} = f(x) = x^T w \tag{18}$$

Where the optimization objective, is to find the best values for the network weights w, that together with the neurons, brings the behavior of the ANN closer to what is expected; which as explained in [25], can be expressed as shown by (19).

$$\arg\min_{\hat{w} \in \mathbb{R}^M} E\left[\|y - \hat{w}^T x\|^2\right] \tag{19}$$

Subsequently, after presenting the basic elements for the ANN to be designed for the controller of the Furuta Pendulum, the network training will be presented below, based on the data acquired from the experimental testbed.

5.1 MRAC-PID through ANN

The original MRAC-PID controller from [11], implemented in the experimental testbed shown in Fig. 6, was used to generate the dataset of the error measurements obtained during the settling process of the pendulum.

Hence, a total of 11,330 values of the pendulum's angle error were acquired, which are used as input values for the ANN. Simultaneously, the data from the motor's duty cycle through time was also acquired, since the duty cycle for the actuator represents the response of the controller (control signal) against the error signal. Thus, output of the network is taken as the control signal for the ANN.

Moreover, the relation between these two variables and their behavior is shown in Fig. 11, where it can be observed that the ideal behavior of the model would be to reduce the duty cycle of the motor when the pendulum is located at 180°.

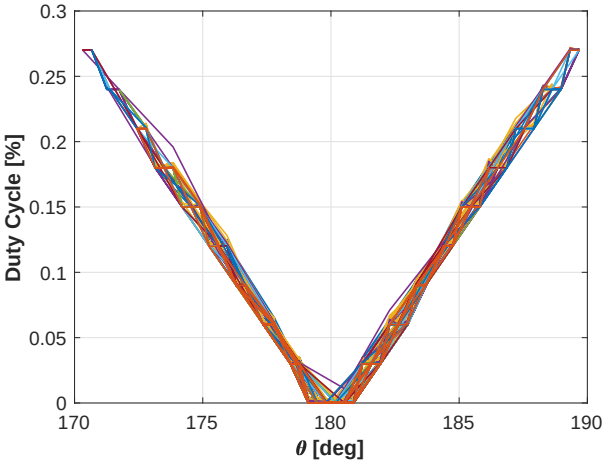


Figure 11: Acquired training data.

Both values (error and duty cycle) were normalized and separated into training, validation, and testing sets, with the 70%, 15%, and 15% relation of the data respectively. The ANN was initialized as described in Table 2, with a MSE loss function and an Adam optimizer with $\alpha = 0.0001$ inspired by [26].

Table 2: ANN architecture

Layer	*Activation func.*	*Neurons*
Input	ReLU	6
Output	Sigmoid	1

Consequently, the network was trained using 100 epochs with a 128 batch size. After the network was trained, the weight matrix was used to evaluate arbitrary error readings in order to generate the regression function, which later was implemented into the embedded system.

Nevertheless, before the experimental test, the resulting matrix was used in order to evaluate the error in a range between 170° and 190°, which is the operating range for the controller in the testbed, seeking to simulate the the resulting duty cycle that would be programmed into the microcontroller. Therefore, the simulated predictions of the control signal against the error signal are shown in Fig. 12.

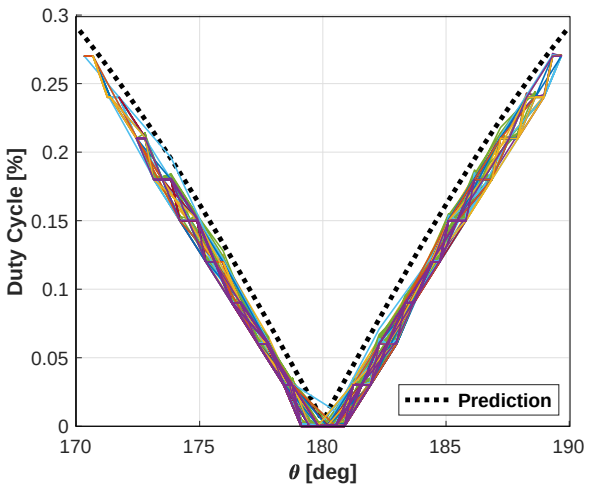


Figure 12: Acquired training data compared to the ANN prediction.

The simulated behavior of the trained network architecture, achieved a similar behavior than the real MRAC-PID controller, with a slightly more aggressive response against the error signal. Which allows predicting a faster response by the ANN-based controller, with a greater performance specially when the pendulum reaches the limit inclination at 170° and 190°.

In addition, the behavior of the predicted results from Fig. 12, demonstrated how the use of an ANN allows more freedom in the system, since the behavior of the net achieves a more regular response of the system, instead of constant changes produced by the modification of the controller coefficients.

Besides, it is clear that in simulation the main advantages of the ANN are not clearly demonstrated, due to the fact that the processing reduction is not evaluated. Thus, it is important to consider and highlight the fact that the obtained prediction, was with simpler arithmetical operations compared to the MRAC-PID controller model; which in a Microcontroller, is traduced into a faster response and into a greater reliability regarding the correct sampling of the error signal, due to the fact that there would be no loss of information or lag in the measurement caused by an improper use of the ticks in the embedded system.

6 Results

The results were achieved through the implementation of the ANN in the embedded system, where one of the main advantages of the proposal is the decrease of the microcontroller's computational effort, when evaluating real-time samples, due the reduction of computational operators through the the simpler mathematical expressions of the ANN.

In other words, it is faster for the hardware to evaluate just a single matrix operation than the three equations that rule the behavior of the coefficients of the PID controller. Such assumption is observed in Fig. 13 and Fig. 14, where the pendulum's angle and system's duty cycle are compared between the original MRAC-PID controller and the ANN model.

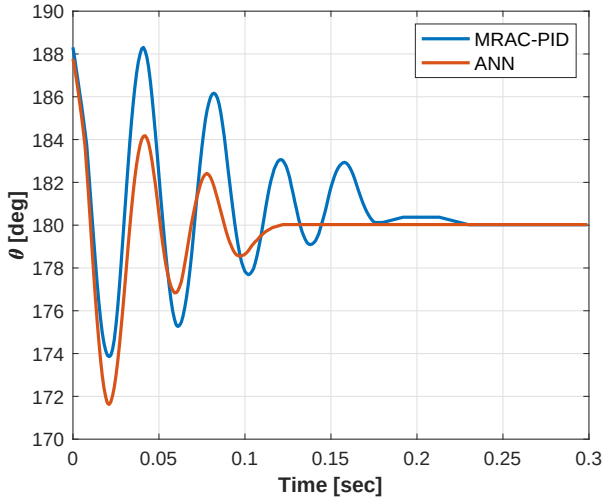


Figure 13: Pendulum's position through time.

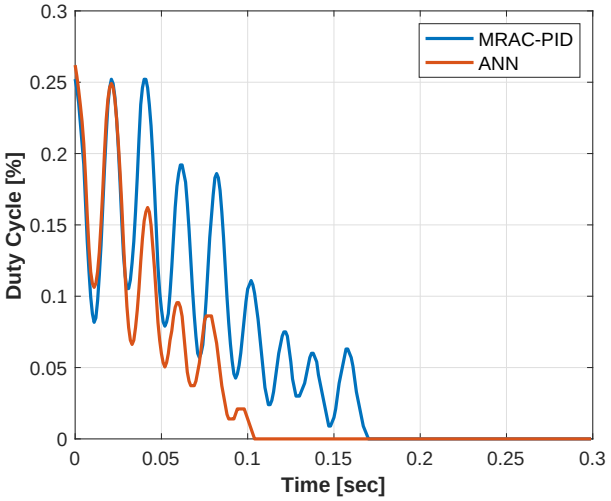


Figure 14: Control signal through time.

Additionally, Fig. 13 shows that the Furuta Pendulum achieves a reduction of the transitional stage, improving the establishment time. As can be clearly seen also in Fig. 13, the controllers were implemented in order to maintain the pendulum in the upright position, therefore the controller begins its operation in the established range (180 ± 10°). Thus, both controllers begin to operate below the 190° for the analysis, stipulated in order to perform fair trails for both controllers.

Therefore, in order to quantify the improvement of the pendulum's behavior, the error between the reference angle (180°) and the actual angle through time, can be measured by performance indexes as shown in [27] and explored by [28]; from where, Integral of Time multiplied by Squared Error (ITSE) and Integral of time multiplied by Absolute Error (ITAE) were the selected indexes.

The first one, is associated to the error energy it is weighted by time, giving more importance to the most recent errors, and the second one related to the absolute error value in time, giving more importance to the errors in the steady-state. Table 4, summarizes the error indexes obtained after evaluating results from Fig. 13.

Table 3: Error indexes of the controllers.

	MRAC-PID	ANN
ITSE	2.3942	1.4032
ITAE	0.5381	0.3114

Hence, by taking the ITSE performance index, the system achieved a 58.61% of error reduction, meanwhile a 57.87% of error reduction was achieved respect to the ITAE index. Then, averaging both results it can be concluded consistently that a 58.24% error reduction was achieved.

On the other hand, last but not the least, Fig. 14 allows to validate the control signal of the controllers, where the [%] of duty cycle through time is shown. Additionally, it can be seen the signal consistency, that is traduced into the reduction of the control effort in time.

The effort reduction can be observed by the way in which the oscillations due to the change in time are compacted, that the ANN controller responds lightly faster than the MRAC-PID, which is enough to allow the faster correction of the pendulum's angle shown in 13. By integrating the response of the controller through time, in can be quantified the reduction in the control effort, achieving then a 58.69% of control effort reduction from the values summarized in Table 4.

Table 4: Control effort.

MRAC-PID	ANN
17.8285	10.4642

Meanwhile, in Fig. 14 it can be analyzed with more detail the settling time improvement, where the system passes from the $T_s \approx 0.17[sec]$ of the MRAC-PID controller, into the $T_s \approx 0.1[sec]$ settling time obtained by the proposed solution.

In spite of the consistency of the results, the authors suggest that another possible implementation to evaluate, could be the use of a Recursive Neural Network (RNN) to evaluate the data as a complete sequential variable, instead of evaluating one measurement at a time; however, the RNN case could result in an equal or more complex model than the original PID controller, which could significantly affect the speed in which the controller responds to the error signal.

Nevertheless, the quantified results from the behavior of the controllers in terms of the angle of the pendulum and the controller effort, validated that the proposed controller achieved a better response as expected, due to the faster response of the system and the computational cost reduction.

Finally, it is important to highlight that the duty cycle signal from Fig. 14, it shown in an absolute value due physical limitations. In real applications such as this one, drivers like the *MD30C*

DC motors driver from the designed testbed, receives two control signals: the first one is the PWM whose modulation leads to the amplitude of the control signal, and the second one is the direction for the motor. Therefore, the direction of the pendulum does not depend on the duty cycle, it is defined through the error's sign.

7 Conclusion

It was observed from the quantification of the results that, not also the behavior of the system followed the original controller's behavior, but also added a more aggressive response when the pendulum was reaching a limit angle that could result in losing equilibrium.

Additionally, results show how a simple neural network model can be narrowed into a simpler regression model, capable of executing similarly as an original and complex controller; such as the MRAC-PID. Nevertheless, the nature of ANN allowed to achieve the stabilization of the pendulum in the desired state, but through less computational effort from the Microcontroller.

The obtained results showed a 58.61% of error reduction respect to the ITSE, a 57.87% reduction respect the ITAE, and a control effort reduction of 58.69%; those three results remarkably allow to conclude that the implementation of the ANN as substitute for the MRAC-PID controller, achieved a better response for the proposed testbed.

Therefore, this works shows an approach methodology for UMS that could allow to improve other designs, due the faster response of the system and its clear lower computational cost from its mathematical expression. Thus, this work enables the method for simpler implementation or even improvements for existing models through artificial intelligence.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment This research is a product of the Project 266632 "Laboratorio Binacional para la Gestión Inteligente de la Sustentabilidad Energética y la Formación Tecnológica" ("Bi-National Laboratory on Smart Sustainable Energy Management and Technology Training"), funded by the *CONACYT* (Consejo Nacional de Ciencia y Tecnologia) SENER (Secretaria de Energia) Fund for Energy Sustainability (Agreement S0019201401).

References

[1] Emese Gincsaine Szadeczky-Kardoss Daniel Szabo. "Novel Cost Function based Motion-planning Method for Robotic Manipulators". In: *Advances in Science, Technology and Engineering Systems Journal* 4.6 (2019), pp. 386–396. DOI: 10.25046/aj040649.

[2] Yang Liu and Hongnian Yu. "A survey of underactuated mechanical systems". In: *IET Control Theory & Applications* 7.7 (2013), pp. 921–935.

[3] JA Acosta. "Furuta's pendulum: a conservative nonlinear model for theory and practise". In: *Mathematical Problems in Engineering* 2010 (2010).

[4] Karl Johan Åström and Katsuhisa Furuta. "Swinging up a pendulum by energy control". In: *Automatica* 36.2 (2000), pp. 287–295.

[5] Navin John Mathew, K Koteswara Rao, and N Sivakumaran. "Swing up and stabilization control of a rotary inverted pendulum". In: *IFAC Proceedings Volumes* 46.32 (2013), pp. 654–659.

[6] Ali Wadi, Jin-Hyuk Lee, and Lotfi Romdhane. "Nonlinear sliding mode control of the Furuta pendulum". In: *2018 11th International Symposium on Mechatronics and its Applications (ISMA)*. IEEE. 2018, pp. 1–5.

[7] Saqib Irfan et al. "Advanced sliding mode control techniques for inverted pendulum: Modelling and simulation". In: *Engineering science and technology, an international journal* 21.4 (2018), pp. 753–759.

[8] Asif Chalanga, Shyam Kamal, and B Bandyopadhyay. "Continuous integral sliding mode control: A chattering free approach". In: *2013 IEEE International Symposium on Industrial Electronics*. IEEE. 2013, pp. 1–6.

[9] Tamen Thapa Sarkar and Lillie Dewan. "Application of LQR and MRAC for swing up control of Inverted Pendulum". In: *2017 4th International Conference on Power, Control & Embedded Systems (ICPCES)*. IEEE. 2017, pp. 1–6.

[10] Adriana Vargas-Martínez and Luis E Garza-Castañón. "Combining adaptive with artificial intelligence and nonlinear methods for fault tolerant control". In: *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer. 2010, pp. 31–41.

[11] Efrain Mendez-Flores et al. "Design, Implementation and Nonlinear Control Analysis of a Furuta Pendulum System". In: *2019 4th International Conference on Control and Robotics Engineering (ICCRE)*. IEEE. 2019, pp. 65–69.

[12] Seyed Hassan Zabihifar, Arkady Semenovich Yushchenko, and Hamed Navvabi. "Robust control based on adaptive neural network for Rotary inverted pendulum with oscillation compensation". In: *Neural Computing and Applications* (2020), pp. 1–13.

[13] Efrain Mendez et al. "Mobile Phone Usage Detection by ANN Trained with a Metaheuristic Algorithm". In: *Sensors* 19.14 (2019), p. 3110.

[14] Slávka Jadlovskỳ and Ján Sarnovskỳ. "Modelling of classical and rotary inverted pendulum systems–a generalized approach". In: *Journal of Electrical Engineering* 64.1 (2013), pp. 12–19.

[15] Daniel Galan et al. "Customized Online Laboratory Experiments: A General Tool and Its Application to the Furuta Inverted Pendulum [Focus on Education]". In: *IEEE Control Systems Magazine* 39.5 (2019), pp. 75–87.

[16] Pavol Seman, Martin Juh, Michal Salaj, et al. "Swinging up the Furuta pendulum and its stabilization via model predictive control". In: *journal of Electrical Engineering* 64.3 (2013), pp. 152–158.

[17] Yun Feng Wu, Zhu Ming, and Ke Chang Fu. "Feedforward and Feedback Control of an Inverted Pendulum". In: *Advanced Materials Research*. Vol. 328. Trans Tech Publ. 2011, pp. 2194–2197.

[18] Mayra Antonio-Cruz et al. "Modeling, simulation, and construction of a furuta pendulum test-bed". In: *2015 International Conference on Electronics, Communications and Computers (CONIELECOMP)*. IEEE. 2015, pp. 72–79.

[19] Ai Xiong and Yongkun Fan. "Application of a PID Controller using MRAC Techniques for Control of the DC Electromotor Drive". In: *2007 International Conference on Mechatronics and Automation*. IEEE. 2007, pp. 2616–2621.

[20] Warren S McCulloch and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity". In: *The bulletin of mathematical biophysics* 5.4 (1943), pp. 115–133.

[21] Aurélien Géron. *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*. " O'Reilly Media, Inc.", 2017.

[22] Thomas Kailath, Ali H Sayed, and Babak Hassibi. *Linear estimation*. BOOK. Prentice Hall, 2000.

[23] Bernard C Levy. *Principles of signal detection and parameter estimation*. Springer Science & Business Media, 2008.

[24] Ali H Sayed. *Fundamentals of adaptive filtering*. John Wiley & Sons, 2003.

[25] Sun Yuan Kung. *Kernel methods and machine learning*. Cambridge University Press, 2014.

[26] Diederik P Kingma and Jimmy Ba. "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980* (2014).

[27] Manuel A Duarte-Mermoud and Rodrigo A Prieto. "Performance index for quality response of dynamical systems". In: *ISA transactions* 43.1 (2004), pp. 133–151.

[28] E Mendez et al. "Electric machines control optimization by a novel geo-inspired earthquake metaheuristic algorithm". In: *2018 Nanotechnology for Instrumentation and Measurement (NANOfIM)*. IEEE. 2018, pp. 1–6.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 325-333 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Performance of a Thermoelectric Powered by Solar Panel for a Large Cooler Box

Mirmanto Mirmanto[1, *], Syahrul Syahrul[1], Made Wirawan[1], I Made Adi Sayoga[1], Agung Tri Wijayanta[2], Ilham Mahyudin[1]

[1]*Jurusan Teknik Mesin, Universitas Mataram, 83125, Indonesia*

[2]*Jurusan Teknik Mesin, Universitas Sebelas Maret, Surakarta, 57126, Indonesia*

ARTICLE INFO

Article history:
Received: 02 November, 2019
Accepted: 19 January, 2020
Online: 07 February, 2020

Keywords:
Thermoelectric
Cooler box
Solar panel
COP

ABSTRACT

An experiment to determine the performance of a thermoelectric powered by solar panel for a large cooler box was carried out. The size of the cooler box tested was 1000 mm x 500 mm x 400 mm and inside the cooler box, a plastic bottle containing 19 liters of water was placed. The thermoelectric hot side was cooled using a mini channel flowed with water. Meanwhile, the cold side of the thermoelectric was connected to the inner heat sink to absorb heat inside the cooler box. The thermoelectric was powered from a battery charged by solar panels directly. There were 2 solar panels, each of which had a power of 100 WP. The cooler box was tested for 6 days in June 2019. The results showed that the lowest temperature on the thermoelectric cold side was around 16.21°C, while the lowest temperature inside the cooler box was 24.25°C. The experimental COPs obtained were ranging from 0.01 to 0.76. Moreover, in general, solar panels were potential as power sources for thermoelectric cooling systems.

1. Introduction

Portable refrigerators are sometimes very necessary, especially in a place far from home or for transportation of goods, blood, and vaccines. For the purposes, thermoelectric cooling machines can be an option because thermoelectric cooling machines are more compact, with no leakage problems, lighter, portable, easier to maintain and more durable as reported in [1- 6]. Nevertheless, the weakness of the thermoelectric is low COP, even less than 1.

There are several ways to raise the COP such as improving the quality of thermoelectric materials, varying the use of thermoelectric techniques, and engineering the dissipation units. The research on thermoelectric materials also has been increased to obtain material that can generate higher COP. But in the term of materials, the thermoelectric development is slow, therefore, attempts to increase the COP through applications have been also tried, one of which is the application to cool beverage cans as reported in [7].

In [7], the authors elucidated that thermoelectric applications to cool beverage cans commonly produced temperatures of 6-8°C within 2 hours. However, the COP of this cooling system was still less than 1, which was 0.856 and could not reach a temperature of 0°C. The authors in [8] investigated the application of thermoelectric modules as a beverage cooler box. They used 2 and 3 thermoelectric modules with a fan or without a fan and a capacity of 34 liters. The beverage cooler box was run for 150 minutes and they obtained a minimum temperature of 14.3°C without water and 16.4°C with 1 liter of water.

In [9], the author examined a thermoelectric refrigerator with 3 thermoelectric modules and power of 51.27 W. The temperature of the refrigerator reached of around 14°C with a drinking water of 1500 ml, while the COP obtained was 0.25. However, even from the engineering side, the application of thermoelectrics is still difficult to obtain high COP. Therefore, this study tries to use a heat dissipation unit that is different from previous studies.

In general, research on the use of thermoelectric cooling modules was carried out using a fin-fan heat sink heat dissipation unit. In [4], authors conducted a study of two different heat dissipation units namely the fin-fan heat sink and the double fan heat pipe. They concluded that with a double fan heat pipe, the cooling box performance including the COP increased. However, the power used by the double fan heat pipe also increased. If the electrical energy consumed by the double fan was taken into account, the heat sink fin-fan unit was superior. Therefore, it is

[*]M. Mirmanto, Jl. Majapahit, no. 62, Mataram, +6282111738971, Email: m.mirmanto@unram.ac.id

necessary to look for other heat dissipation units such as water heat exchangers.

The water heat exchangers such as a water block used for removing heat from the hot side of thermoelectrics had been studied in [10]. Apparently, the water block was still not effective because the ability to transfer heat was still low. Water heat exchangers that might be able to increase COP are mini-channel or mini-pin fin. Why a mini channel is predicted to be able to increase COP? Based on the previous research, mini-channel has been presenting to have a higher heat transfer ability compared to other water heat exchangers as reported in [11]. Similarly, the mini pin fin heat exchanger had been tested to be able to transfer more heat as explained in [12].

Furthermore, COP is also influenced by the size of the refrigerator. A large refrigerator increases the COP and the temperature. Nevertheless, the high refrigerator temperature means that the cooling process is not adequate. Likewise, vice versa, a small refrigerator produces a small COP and a low temperature. However, the effect of refrigerator size is unclear for the refrigerator with a large mass of goods. The more mass in the refrigerator, the higher COP and temperatures as explained in [5]. Table 1 shows several sizes of cooler boxes that have been studied. In Table 1, the largest refrigerator/cooler box size studied is 620 mm x 595 mm x 1565 mm. The refrigerator was examined in [15]. However, they used 8 thermoelectric modules.

Power for thermoelectric refrigerators is generally obtained from conventional electric grids. However, in Indonesia, the price of electricity is still relatively high and the availability of electricity has not been fully distributed to remote areas and small islands yet. Therefore, this study uses electricity sources obtained from solar panels. Even so, this paper does not discuss solar panels in detail because the solar panels are only used as the DC power source.

This research has the advantage of overcoming the urgent problem: (i) the supply of fossil fuels is almost gone, so all electrical equipment including refrigerators may be operated with renewable energy, (ii) introducing new cooling methods that do not require large and expensive equipment, (iii) thermoelectric cooler boxes with mini-channel heat dissipation units have not been widely studied. The purpose of this study is to examine the performance of a large refrigerator using a single thermoelectric module powered using solar panels. The thermoelectric hot side is cooled utilizing a water mini-channel. The characteristics of the refrigerator are represented by the lowest temperatures that can be achieved and the COP. The contribution of this research to the thermoelectric study field is to increase the information concerning thermoelectric usages for cooler boxes.

2. Research Method

2.1. Experimental Set-up

This work was experimental research using a test rig as shown in Figure 1 consisting of a cooler, a thermoelectric, solar panels, water, a mini-channel, a pump, a small radiator, a small tank, a flow meter, a battery, a solar charger, and a data logger National Instrument (NI9714). All temperatures were measured using K type thermocouples calibrated with an accuracy of ± 0.5°C. The electrical power used was measured using a Vichy Vc8145 digital multimeter, while the flow of the water passing through the mini-channel was measured using an FLR1000 flowmeter with an accuracy of ± 0.2 g/s.

When the electric current was supplied to the thermoelectric, the thermoelectric hot side temperature increased drastically. This increased temperature should be maintained so that it did not exceed 40°C, i.e. using water flow. Finally, the heat from the thermoelectric was carried away by the flow of water towards the small radiator. Inside the small radiator, the heat was discharged into the environment through the radiator walls. Then the water returned to the small tank. The electricity used fluctuated because the source was solar panels. When the sun was bright, the power was great and vice versa. However, these fluctuations could be prevented using a battery.

Table 1: Sizes of cooler boxes that had ever been studied, heat dissipation units, and COPs

Reference	Size	Heat dissipation unit	COP
[1]	180 mm x 230 mm x 320 mm	Heat sink with a fan	0.16
[3]	-	Heat sink with a fan	Decrease with difference temperature
[4]	285 mm x 245 mm x 200 mm	Heat sink with a fan, and double fan heat pipe	0.002-0.02
[7]	16 mm x 8 mm x 9 mm	Heat sink with a fan	0.856; 0; 731
[8]	500 mm x 315 mm x 291 mm	Heat sink with a fan	-
[13]	-	Heat sink with a fan	0.43-0.45
[14]	250 mm x 250 mm x 350 mm	Heat sink with a fan	0.65
[15]	620 mm x 595 mm x 1565 mm	Heat sink, big aluminum	0.58
[16]	1. 65 mm x 70 mm x120 mm. 2. 55 mm x 57 mm x 140 mm 3. 100 mm x 100 mm x 230 mm 4. 60 mm x 65 mm 150 mm 5. 80 mm x 75 mm x 170 mm	Heat sink with a fan	-
[17]	-	Heat sink with fan	0.03-0.194

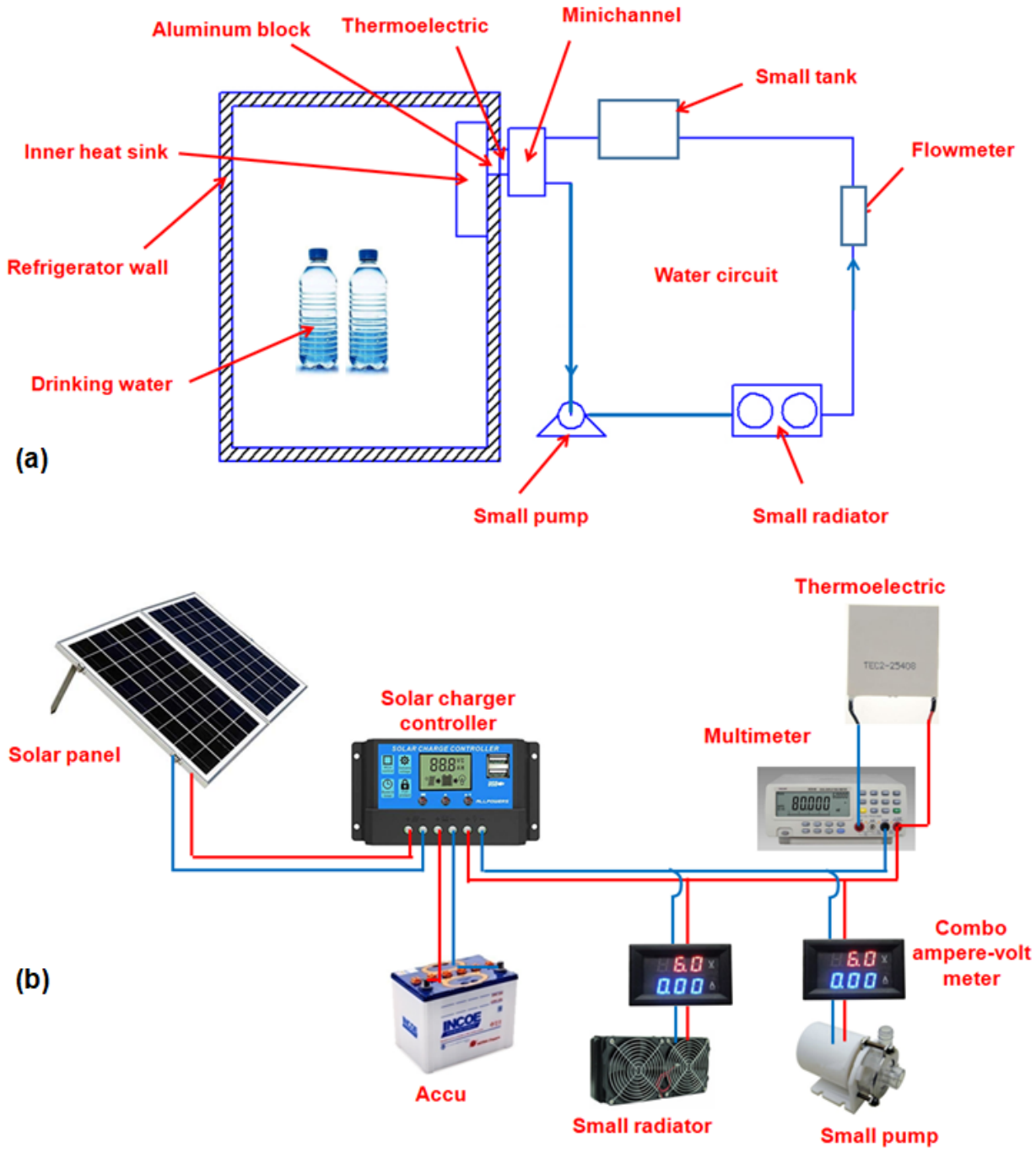


Figure 1: Schematic diagram of the research apparatus; (a) refrigerator and water circuit, (b) electrical circuit.

Thermocouple placements can be seen in Figure 2. The thermocouple locations are denoted using letters a-q. The solar panels used were two pieces with each power of 100WP, and the battery employed was a dry (free maintenance) battery with a voltage of 12 V and a current of 100 AH. However, the analysis here is focused on the performance of the refrigerator only and deep analysis of the mini-channel is not given in this paper.

2.2. Heat Transfer Analysis

The experimental data were analyzed using various equations to determine the performance of the refrigerator. All parameters of the research were analyzed to obtain conclusions. The heat balance in the thermoelectric can be estimated using equation (1) which can be obtained in [14].

$$Q_h = Q_c + P_{in} \tag{1}$$

Q_h is the heat released on the hot side of the thermoelectric (W), Q_c is the heat absorbed on the cold side of the thermoelectric (W) and P_{in} is the input power (W).

In the design of thermoelectric refrigerators, there were two types of heat loads that should be overcome, namely (i) the heat load from the air and the items in the refrigerator, and (ii) the heat

load entering through the walls of the cooler box. The equation to determine the heat load from air and goods can be seen in [4-6, 18-19], which can be expressed in the equation as follows:

$$Q = \frac{\sum_{1}^{n} m_{(i)} c_{p(i)} \left(T_i - T_{i-1}\right)}{t} \quad (2)$$

Q is the heat load and is defined as the total heat coming from the goods inside the cooler box (W). $m_{(i)}$ is the momentary mass (kg), $c_{p(i)}$ is the momentary specific heat of the goods (J/kg K) and $T_{(i)}$ represents the momentary temperature (°C), while t is the total time of the running machine (s). Whereas heat entering through the refrigerator wall that is called conduction can be estimated by the equation that can be found in [20].

$$Q_k = -kA \frac{\sum_{1}^{n} dT_{(i)} t_{(i)}}{dxt} = -\frac{\Delta T_{(i)} t_{(i)}}{Rt} \quad (3)$$

Q_k is the conduction heat transfer rate (W), k is thermal conductivity (W/m°C), A denotes heat transfer area (m²), and $dT_{(i)}/dx$ is the instantaneous temperature gradient over a distance (°C/m). $\Delta T_{(i)}$ is the instantaneous temperature difference of the hot and cold surface temperatures (°C), while R is a thermal resistance (°C/W). Total cooling capacity can be determined by the equation that can be obtained in [18].

$$Q_c = Q_k + Q \quad (4)$$

Q_c is the total cooling load (W), and the COP of the refrigerator can be computed using equation (5) that can be attained in [21]:

$$COP = \frac{Q_c}{P_{in}} = \frac{Q_c}{VI} \quad (5)$$

The P_{in} is all power given to the refrigerator system including the power given to the thermoelectric, to the pump, and to the small radiator. V is the voltage, and I is the current. The thermoelectric hot side was cooled using a mini-channel, in which water flowed. To flow the water, a mini pump was installed on the system, and the pump required electrical power as well. In addition to the pump, the small radiator also required power, because this small radiator utilized two fans, which were driven by DC electricity. So the total power used was the sum of the thermoelectric power, pump power, and small radiator power. COP was the comparison of Q_C to total power as shown in equation (5). All power was taken from a battery, and the battery was filled by solar panels.

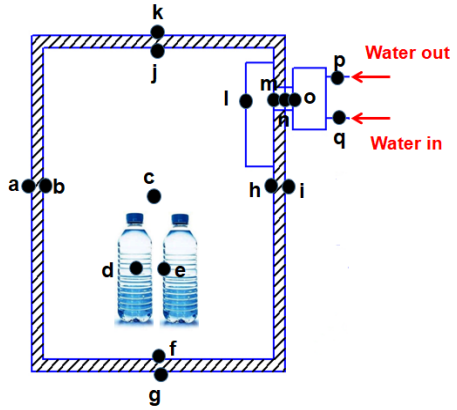


Figure 2: Thermocouple locations are marked with the letter a-q

3. Results and Discussion

The results of the study were presented in the form of graphs and tables. Before the results of the experiments were presented, the recorded temperature results were given in Figure 3 for 6 days. The purpose of presenting this figure was to see the temperature trend and the temperature differences for 6 days.

The thermoelectric hot side temperature increased drastically and then flattened out. This was because heat, Q_h, came out from the thermoelectric hot side. This was also observed by several previous researchers, e.g. [1, 4-6, 14, 22]. Meanwhile, the thermoelectric cold side temperature dropped dramatically and then flattened out. The thermoelectric cold side temperature trend also showed a similar trend, but the room temperature was higher than the thermoelectric cold side temperature. This was because the heat transferred from the room to the cold side of the thermoelectric went through a long connector. In the connector, there could be other heat coming from the surrounding, and then the heat absorbed in the cooler box room was lower than the heat absorbed by the cold side of the thermoelectric.

From the air in the refrigerator, the heat was absorbed by the inner heat sink. From the inner heat sink, the heat went through an aluminum block to arrive at the surface of the thermoelectric cold side. However, at the same time, there was the heat that entered through the refrigerator wall. The heat might come to the connector as well. As a result, the cold side of the thermoelectric was unable to absorb much heat from the refrigerator because it had gained additional heat at the connector.

From the cooler box temperature, water temperature, and bottle temperature, then the total heat that should be removed from inside the cooler box could be estimated. The estimation of heat transfer from air, water, and bottles can be done using equation (2). The results of the experiment showed that the heat taken from the air, from water, and from the bottle decreased with time, and then flattened out, see Figures 4-7. This was because the heat inside the cooler box could not be absorbed anymore after several seconds. The thermoelectric capability was maximal in these conditions. However, such things were also obtained by previous researchers such as [4-6, 14, 23]. The parameter used to indicate the performance of the cooler box was COP. COP could be determined using equation (5). The COP trend decreased with time. In the initial seconds, the COP was very high, then the COP dropped continuously over time, see Figure 10.

This COP trend was also found by some previous researchers such as [13-14, 22, 24-26]. In [26], the authors drained water on the thermoelectric hot side using a mini-channel. However, they used conventional electricity sources and many thermoelectric modules, while in this experiment, the electrical power was taken from solar panels and the thermoelectric module used was only one. How to flow water in [26] was also different from this study.

They flowed water from the reservoir to the pump, to the thermoelectric hot side, to the heat exchanger (small radiator), and back to the reservoir. In this study, the water was flowed from the reservoir to the thermoelectric hot side, to the pump, to the small radiator and back to the reservoir. So the difference was the location of the pump installed on the system/ test rig.

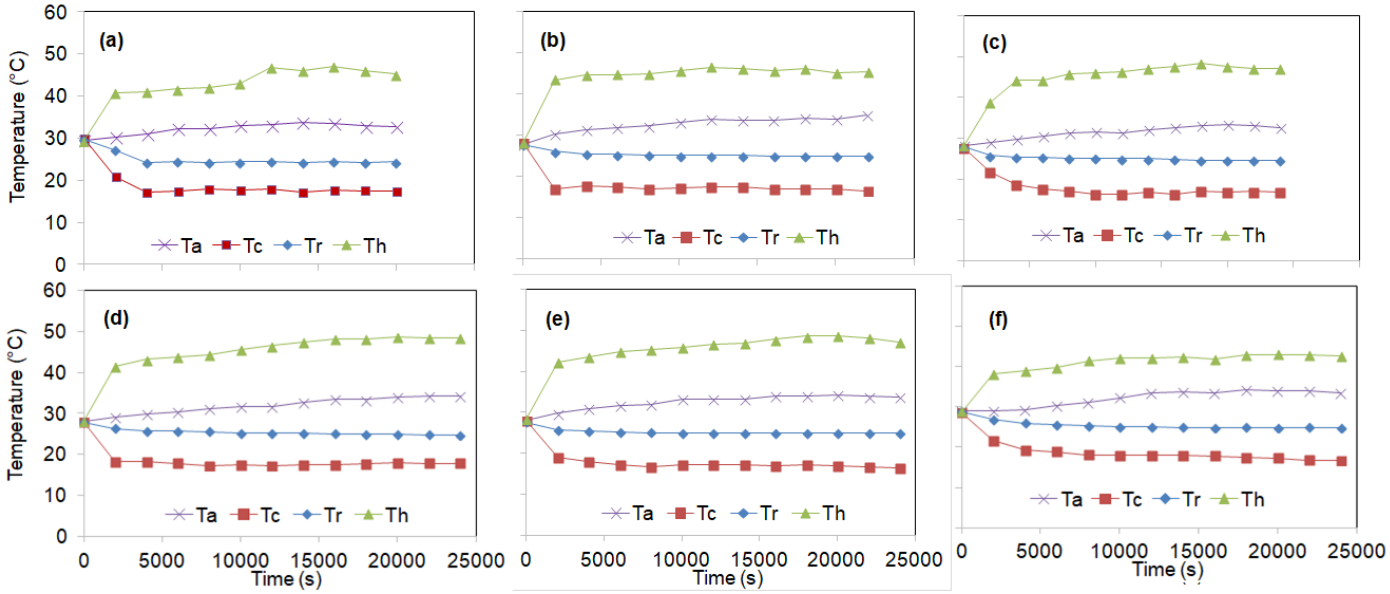


Figure 3: Trend of temperatures with time for 6 days; (a) day 1, (b) day 2, (c) day 3, (d) day 4, (e) day 5, (f) day 6.

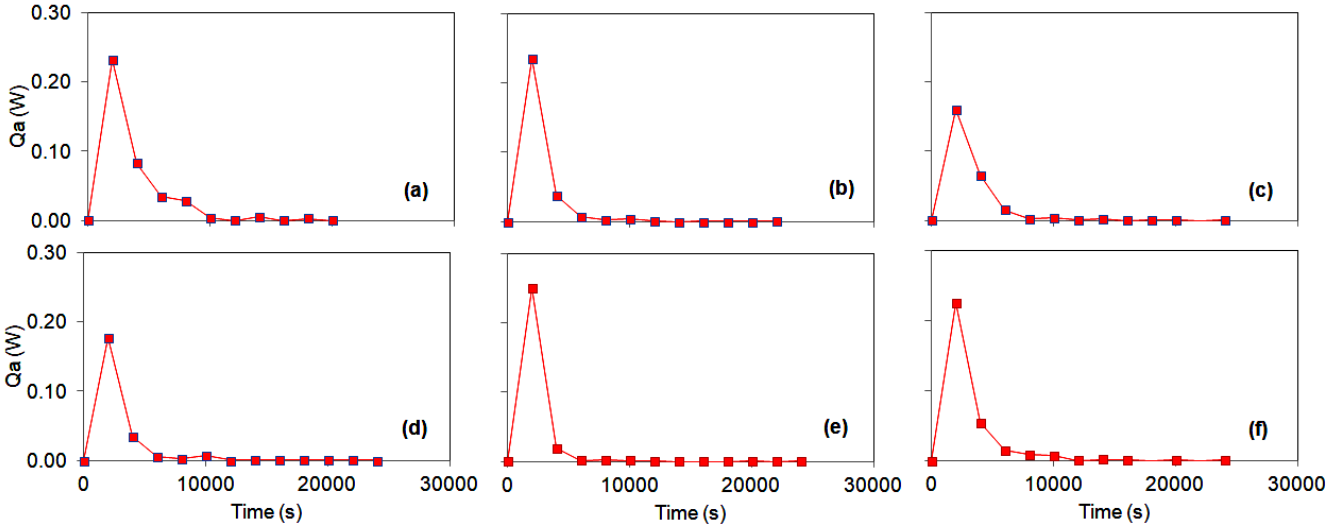


Figure 4: Calculated heat absorbed from the air inside the refrigerator; (a) day 1, (b) day 2, (c) day 3, (d) day 4, (e) day 5, (f) day 6.

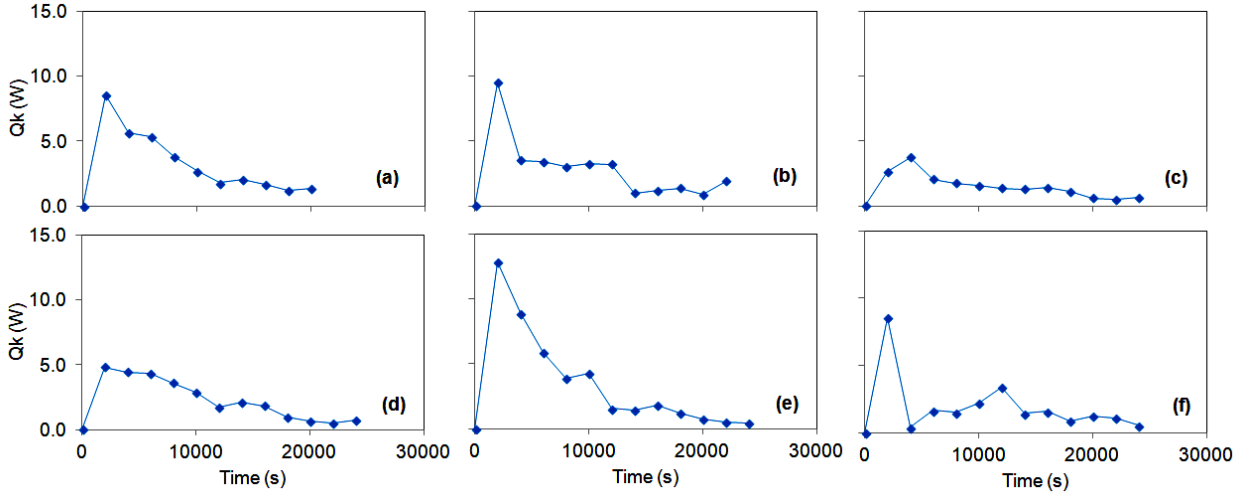


Figure 5: Trends of the conduction heat transfer rate with time; (a) day 1, (b) day 2, (c) day 3, (d) day 4, (e) day 5, and (f) day 6

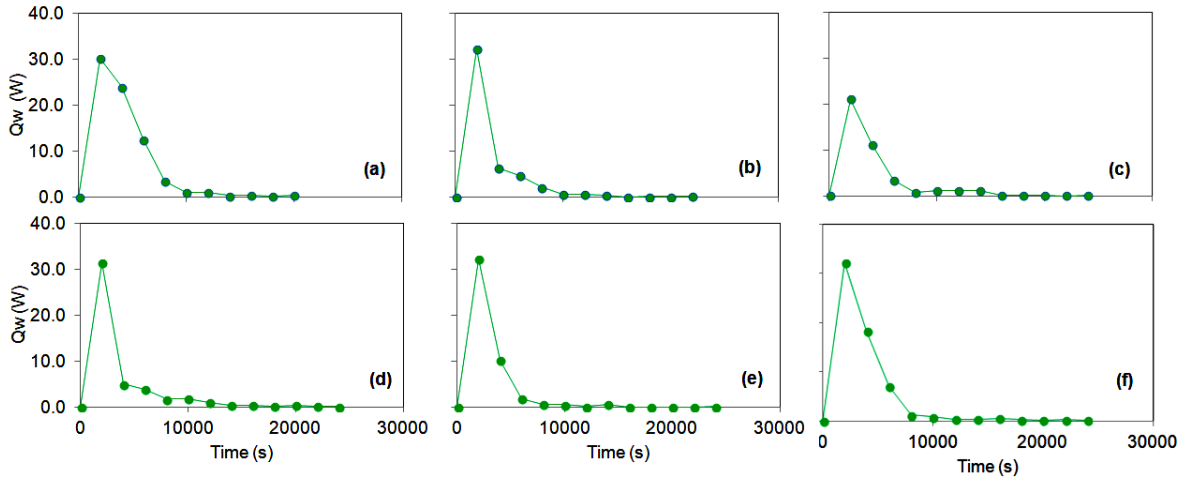


Figure 6: Trends of the heat transfer rate absorbed from the water respect to time; (a) day 1, (b) day 2, (c) day 3, (d) day 4, (e) day 5, and (f) day 6

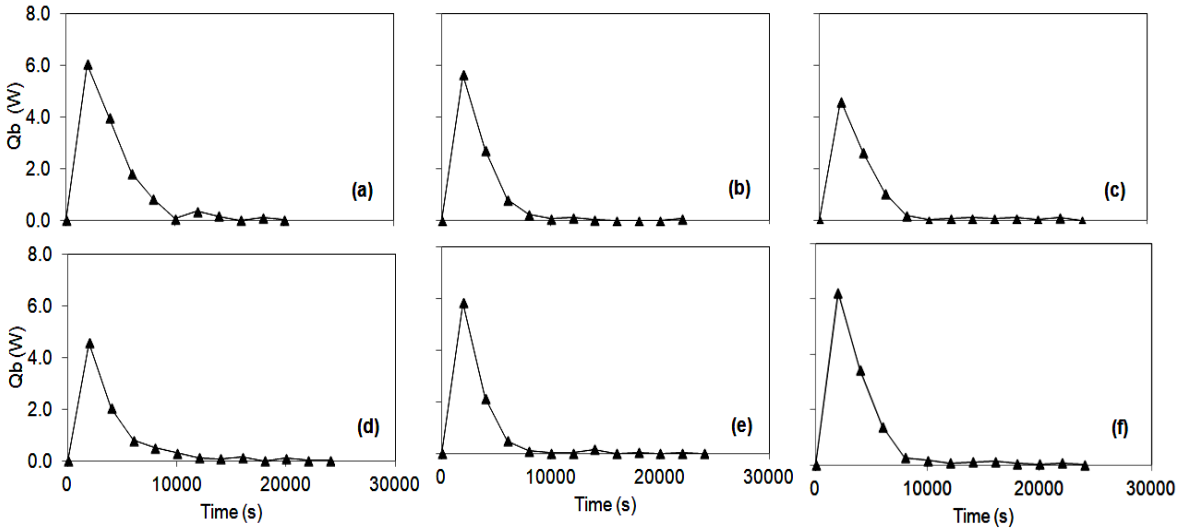


Figure 7: Trends of the heat transfer rate absorbed from the plastic bottle with time; (a) day 1, (b) day 2, (c) day 3, (d) day 4, (e) day 5, and (f) day 6

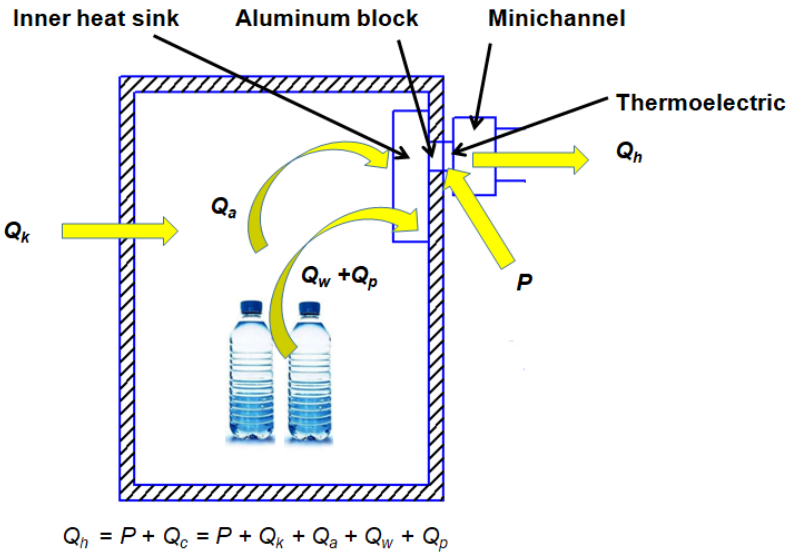


Figure 8: Illustration of heat flows in the refrigerator system for this study

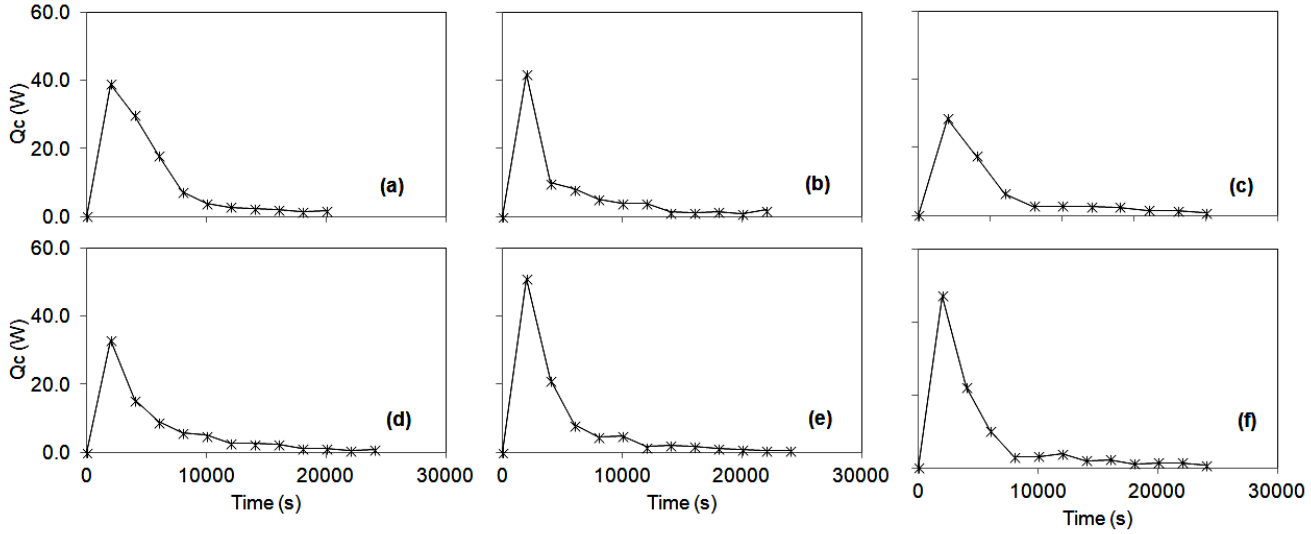


Figure 9: Trends of total heat transfer rate that should be removed from inside the refrigerator with time; (a) day 1, (b) day 2, (c) day 3, (d) day 4, (e) day 5, and (f) day 6

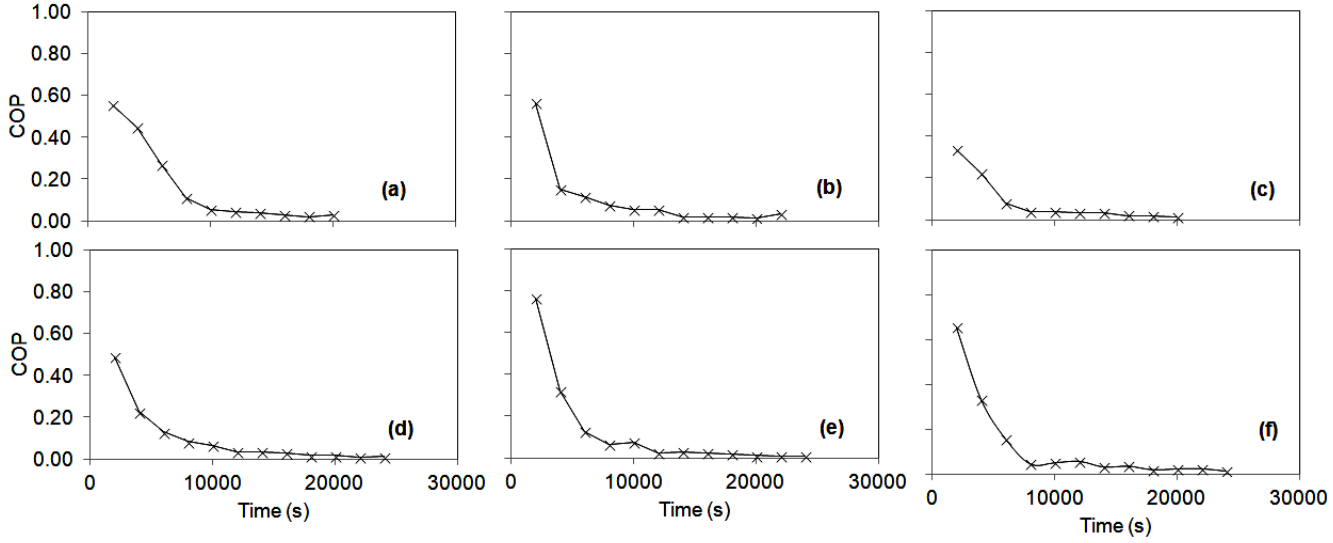


Figure 10: Trends of the COP with time; (a) day 1, (b) day 2, (c) day 3, (d) day 4, (e) day 5, and (f) day 6

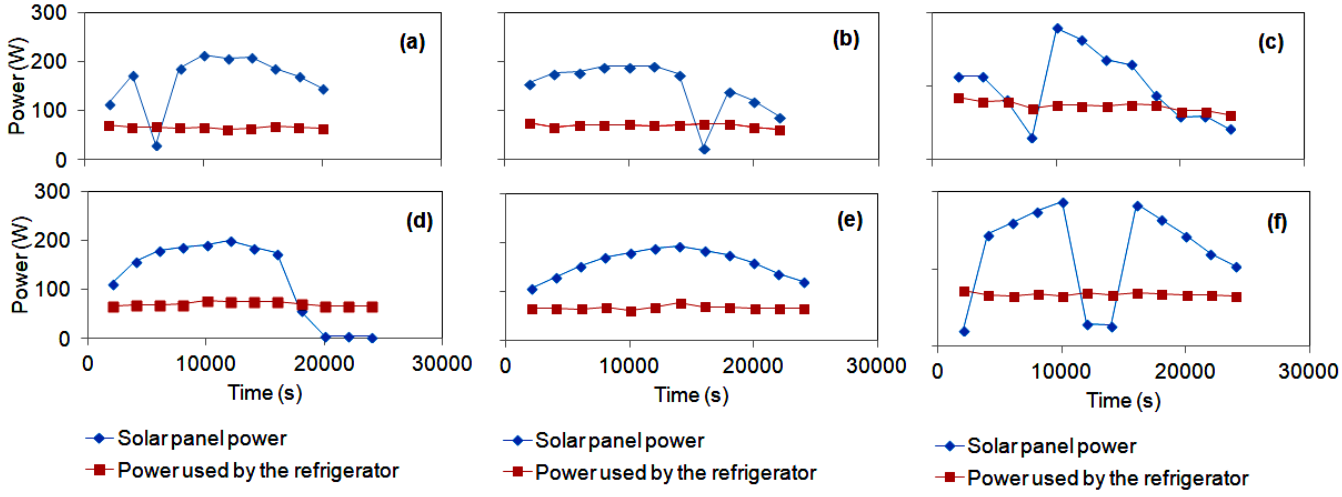


Figure 11: Solar panel power and power used by the refrigerator with time

From the character of the refrigerator mentioned above, the most interesting thing was the relation of the character to the output power of the solar panel. The output power of the solar panels during the study was recorded and shown in Figure 11. The power of solar power increased with time and finally after reaching its peak it dropped again. However, solar panel power was very dependent on sky conditions. If the sky was suddenly cloudy, the solar panel power went down and vice versa. Therefore, the increase in solar panel power was uncertain. Besides that, in this study, the solar panel power was determined as the multiplication of solar power and solar panel area. Furthermore, the multiplication results were then multiplied by the efficiency of solar panels which in general was around 20%. Apparently, fluctuations in solar panel power did not affect the power used by the refrigerator. This was because the output of the solar panel was connected to the battery so that the fluctuations could be muted by the battery. The number of solar panels used was 2 pieces, and each had a power of 100 WP, so the total power was around 200 WP. The power output was apparently less than 200 W as shown in Figure 11.

4. Conclusion

An experimental study to measure the performances of thermoelectric powered using solar panels for cooling a big cooler box was conducted. The cooler box had one thermoelectric module. The air temperature inside the cooler box was still high. One thermoelectric was not adequate for cooling purposes for a big cooler box. The number of thermoelectrics installed should be more than 1 in order to fulfill the cooling purposes. Due to the battery, the fluctuation of the solar panel output became collapse and the power absorbed by the thermoelectric was stable for about 25200 s. In general, the solar panel could be used for powering the thermoelectric cooler box, but one module was not enough for a big box.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to acknowledge the DRPM for the funding through the research PDUPT schema.

Nomenclature

A	cross-sectional area (m²)
c_p	specific heat (J/kg°C)
COP	coefficient of performance
dT	temperature difference (°C)
dx	distance (m)
I	current (A)
m	mass (kg)
Q	heat transfer rate (W)
R	thermal resistance (°C/W)
t	time (s)
T	temperature (°C)
P	power (W)
V	voltage (V)
ΔT	temperature change, temperature difference (°C)

Subscript

a	air
c	cold
h	hot
in	input
k	conduction
p	plastic bottle
w	water

References

[1] S.A. Abdul-Wahab, A. Elkamel, A.M. Al-Damkhi, I.A. Al-Habsi, H.S. Al-Rubai'ey', A.K. Al-Battashi, A.R. Al-Tamimi, K.H. Al-Mamari, M.U. Chutani, "Design and Experimental Investigation of Portable Solar Thermoelectric Refrigerator" Renewable Energy, 34, 30-34, 2009. https://doi.org/10.1016/j.renene.2008.04.026

[2] M.K. Rawat, H. Chattopadhyay, S. Neogi, "A Review on Developments of Thermoelectric Refrigeration and Air Conditioning Systems: A Novel Potential Green Refrigeration and Air Conditioning Technology" Int. J. Engineering Technology and Advanced Engineering, 3(3), 362-367, 2013.

[3] N.J.M. Reddy, "A Low Power, Eco-Friendly Multipurpose Thermoelectric Refrigerator" Front. Energi, 10(1), 79-87, 2016. https://doi.org/10.1007/s11708-015-0380-8

[4] Mirmanto, I.B. Alit, I.M.A. Sayoga, R. Sutanto, Nurchayati, A. Mulyanto, "Experimental Cooler Box Performance Using Two Different Heat Removal Units: A Heat Sink Fin-Fan, and a Double Fan Heat Pipe" Frontier in Heat and Mass Transfer, 10(34), 1-7, 2018. http://dx.doi.org/10.5098/hmt.10.34

[5] M. Mirmanto, I.W. Joniarta, I.M.A. Sayoga, N. Nurpatria, Y.A. Padang, I.G.N.K. Yudhyadi, "Effect of Water Volume on a Thermoelectric Cooler Box Performance" Frontiers in Heat and Mass Transfer, 11(17), 1-9, 2018. http://dx.doi.org/10.5098/hmt.11.17

[6] M. Mirmanto, S. Syahrul, Y. Wirdan, "Experimental performances of a thermoelectric cooler box with thermoelectric position variations" Engineering Science and Technology, an International Journal, 22, 177–184, 2019. https://doi.org/10.1016/j.jestch.2018.09.006

[7] B. Irwin, A. Ruri, "Kaji Eksperimental Kotak Pendingin Minuman Kaleng dengan Termoelektrik Bersumber Dari Arus Dc Kendaraan dalam Rangkaian Seri dan Paralel", *Seminar Nasional Tahunan Teknik XII*, Bandar Lampung, 2013.

[8] A. Aziz, S. Joko, S. Villager, "Aplikasi Modul Pendingin Termoelektrik Sebagai Meia Pendingin Kotak Minuman" Jurnal Teknik Mesin Universitas Riau, 32-38, 2014.

[9] G. Ramdhan, "Sistem Pendinginan pada Kulkas Termoelektrik dengan Variasi Pembebanan" Skripsi, Jurusan Teknik Mesin, Fakultas Teknik, Universitas Mataram, Mataram, NTB, Indonesia, 2016.

[10] D.K. Putra, "Unjuk Kerja Kotak Pendingin Termoelektrik dengan Variasi Laju Aliran Massa Air Pendingi" Skripsi, Jurusan Teknik Mesin, Universitas Mataram, 2018.

[11] J. Hejcik, M. Jicha, "Single-Phase Heat Transfer in Minichannels" EPJ Web of Conference, 67, 1-4, 2014. http://doi.org/10.1051/epjconf/20146702034

[12] T. Yeom, T. Simon, T. Zhang, M. Zhang, M. North, T. Cui, "Enhanced Heat Transfer of Heat Sink Channels With Micro Pin Fin Roughened Walls" Int. J. of Heat and Mass Transfer, 92, 617–627, 2016. https://doi.org/10.1016/j.ijheatmasstransfer.2015.09.014

[13] M. Gillott, L. Jiang, S. Riffat, "An Investigation of Thermoelectric Cooling Devices for Small-Scale Space Conditioning Applications in Buildings" International Journal Of Energy Research, 34, 776-786, 2010. http://doi.org/10.1002/er.1591

[14] S. Jugsujinda, A. Vora-ud, T. Seetawan, "Analysing of thermoelectric refrigerator performances" Procedia Engineering, 8, 154-159, 2011. https://doi.org/10.1016/j.proeng.2011.03.028

[15] K. Atik, Y. Yildiz, "An Experimental Investigation of a Domestic Type Solar TE Cooler" Energy Sources, Part A, 34, 645–653, 2012. http://dx.doi.org/10.1080/15567031003627989

[16] A. Aziz, R.I. Mainil, A.K. Mainil, Syafri, M.F. Syukrillah, "Design of Portable Beverage Cooler Using One Stage Thermoelectric Cooler (TEC) Module" Aceh Int. J. Sci. Technol., 7(1), 29-36, 2017. http://doi.org/10.13170/aijst.6.1.5427

[17] D. Suryawanshi, N. Pokharkar, V. Pokale, A. Walgude, "Design and Fabrication of Thermoelectric Refrigerator for Liquid Cooling by Automatic

Temperature Micro-Controller" International Journal of Science Technology & Engineering, **3**, 21-27, 2016.

[18] F. Incropera, D.P. DeWitt, T.L. Bergman, A.S. Lavine, "Fundamentals of Heat and Mass Transfer, 6th Edition, John Wiley & Sons, Hoboken, 2006.

[19] H.P. Banjarnahor, "Analisis Laju Pendinginan pada Kulkas Thermoelektrik Super Cooler Dibandingkan Sistem Pendingin Konvensional Menggunakan Gas Freon" Skripsi, USU Institutional Repository, Universitas Sumatera Utara, 2016.

[20] A. Waleed et al., "Solar (PV) Water Irrigation System with Wireless Control," in 2019 International Symposium on Recent Advances in Electrical Engineering (RAEE), 2019, vol. 4, pp. 1–4.

[21] Y.A. Cengel, M.A. Boles, "Thermodynamics an Engineering Approach", 6th Ed., McGraw-Hill Higher Education, Boston, 2008.

[22] S.M. Hafis, M.J.M. Ridzuan, A.Z.A. Firdaus, S.M. Shahril, R.N. Farahana, C.A. Chong, "COP Improvement of Thermoelectric Cooler Through the Optimization of Heat Dissipation System" Applied Mechanics and Materials, 554, 241-245, 2014. https://doi.org/10.4028/www.scientific.net/AMM.554.241

[23] H. Ananta, Mirmanto, Y.A. Padang, "Unjuk Kerja Kulkas Termoelektrik dengan Rangkaian Seri dan Parallel pada Beban Air 1500 ml" Dinamika Teknik Mesin, 7(2), 80-86, 2017. https://doi.org/10.29303/dtm.v7i2.157

[24] B.J. Huang, C.J. Chin, C.L. Duang, "A Design Method of Thermoelectric Cooler" International Journal of Refrigeration, 23, 208-218, 2000. https://doi.org/10.1016/S0140-7007(99)00046-8

[25] H. A. Raza, M. T. Riaz, M. Idrees, M. M. Afzal, W. S. Hashmi, and M. W. Asif, "Analysis the effect of 500kv High-Voltage Power Transmission Line on the Output Efficiency of Solar-Panels," in 2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), 2019, pp. 1–6.

[26] M. Gökçek, F. Şahin, "Experimental Performance Investigation of Minichannel Water Cooled-Thermoelectric Refrigerator", Case Studies in Thermal Engineering, 10, 54-62, 2017. https://doi.org/10.1016/j.csite.2017.03.004

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 6, 328-338 (2019)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

A Method for Mosaicking Aerial Images based on Flight Trajectory and the Calculation of Symmetric Transfer Error per Inlier

Daniel Arteaga[*,1], Guillermo Kemper[1], Samuel G. Huamán Bustamante[1], Joel Telles[1], León Bendayán[2], José Sanjurjo[2]

[1] *Universidad Nacional de Ingeniería, Instituto Nacional de Investigación y Capacitación de Telecomunicaciones, Lima, 15021, Perú*

[2] *Instituto de Investigación de la Amazonía Peruana, Iquitos 16007, Perú*

ARTICLE INFO

Article history:
Received: 16 October, 2019
Accepted: 30 November, 2019
Online: 23 December, 2019

Keywords:
Image Mosaicking
Random Sample Consensus
Georeferencing
Unmanned Aerial Vehicles
Forest Monitoring

ABSTRACT

In recent years, development of aerial autonomous systems and cameras have allowed increasing enormously the number of aerial images, and their applications in many research areas. One of the most common applications is the mosaicking of images to improve the analysis by getting representations of larger areas with high spatial resolution. This paper describes a simple method for mosaicking aerial images acquired by unmanned aerial vehicles during programmed flights. The images were acquired in two scenarios: a city and a forest in the Peruvian Amazon, for vegetation monitoring purposes. The proposed method is a modification of the automatic homography estimation method using the RANSAC algorithm. It is based on flight trajectory and the calculation of symmetric transfer error per inliers. This method was implemented in scientific language and the performance was compared with a commercial software with respect to two aspects: processing time and geolocation errors. We obtained similar results in both aspects with a simple method using images for natural resources monitoring. In the best case, the proposed method is 6 minutes 48 seconds faster than the compared software and, the root mean squared error of geolocation in X-axis and Y-axis obtained by proposed method are less than the obtained by the compared software in 0.5268 and 0.5598 meters respectively.

1 Introduction

During most of the twentieth century, the photogrammetry required a lot of monetary and time resources. After the technological development of electronic systems like the Global Positioning System (GPS) and the Inertial Measurement Unit (IMU), the design of autonomous vehicles, and the production of digital cameras, time and cost limitations have been steadily declining. Thus, the rise of the Unmanned Aerial Vehicles (UAVs) started. Nowadays, there is a wide range of UAVs with multiple sensors and digital cameras that are being used in many different areas such as photogrammetry, agriculture, management of natural resources, mapping and urban planning, rescue [1], assessment and mitigation of disasters, and other Remote Sensing applications [2].

One of the most important and common uses given to the aerial images acquired by UAVs, also known as drones, is the image mosaicking, which allows better visualization and assessment of an event or area of study. For instance, the usage of image mosaicking in agriculture [3] is beneficial not just because it generates an image that covers bigger areas of crops with high resolution, in fact, it also allows control the temporal resolution and the periodical assessment of the fields to improve decision-making regarding the amounts of pesticides, irrigation, land usage, and other variables.

Image mosaicking is about the reconstruction of a scene in two dimensions (2D) by using individual images with overlapping areas. In order to perform this process, the estimation of geometric transformations between pairs of images is required to project them over the others and merge them together with the minimum possible error. These projective transformations, which are commonly known as 2D homographies, are estimated from matching points of two overlapping images.

Multiple mosaicking algorithms have been developed depending on the requirements of accuracy and time. There are high accuracy algorithms based on Structure from Motion technique that estimates the scene information in three dimensions, the orientation and the localization of the camera, and in this way, it calculates the homographies [4]; however, these algorithms imply high computational cost and are difficult to implement. Other algorithms utilize the information of the GPS and gyroscope of the UAV to obtain approximations of localization and orientation of the cameras during

[*]Daniel Arteaga, Lima, 15067, Peru, (+51) 995550108 & darteaga@inictel-uni.edu.pe

the flight [5] but matching images and gyroscope information by the acquisition time may not be accurate. In addition, it has been implemented algorithms that perform feature matching based on bag of words (BoW) and dictionaries [6, 7] to accelerate this step, for example, in [8] author compared processing times of different methods of feature matching. Moreover, other researchers have modified the Structure from Motion technique and have changed the Bundle Adjustment by other techniques of global adjustment [9, 10]; nevertheless, these algorithms are difficult to implement.

We present a simple method of mosaicking aerial images with geographical information acquired by UAV that obtains viable results for applications in natural resources monitoring, and competitive results in terms of processing time and geolocation error, in comparison with professional software. This method is based on the image mosaicking of aerial images acquired in the same line of flight, and the usage of symmetric transfer error per inlier as an indicator of acceptable estimation of the homography.

This work is organized as follow: Section 2 develops the proposed mosaicking algorithm and the implementation details. In section 3, the results are presented and compared with those of one professional software. Finally, in section 4, we present the conclusions and comment some limitations of the proposed algorithm.

2 Proposed Method

The proposed method is based on a modification of the Random Sample Consensus (RANSAC) method used to automatically estimate homographies. According to Figure 1, this method first creates a mosaic of all images acquired in the same line of flight and then obtains the final mosaic by merging properly the mosaics of flight lines. Given that our method is based on the flight trajectory, first, we mention some details related to the acquisition of images via UAV: the images must be composed of three channels (the most common case is RGB), and they should be geotagged (latitude and longitude) in order to obtain flight lines and to perform the georeferencing step. In this section, we describe and explain the implementation details regarding the steps of the proposed method.

2.1 Image Reading and Image Preprocessing

The first step is to read geotagged images and obtain the number, the names, the directory, the size and the geographical information of the images. Then the images are resized to a maximum resolution of 1500 x 1000 pixels and the geographical information is transformed to the Universal Transverse Mercator (UTM) coordinate system.

The preprocessing consists of the selection of images that belong to the same flight lines, which are almost straight lines as those of Figure 2. It is important to understand that the geolocation information of each image corresponds to the point located in the middle of the image. To determine which images belong to the same lines we propose an approach based on the distance between two adjacent images also known as shooting distance.

If the shooting distance between two images is inside a certain range of distances, the method will consider those images have been acquired in the same flight line. The range of distances is established from the most common values of shooting distances and it is calculated as follow:

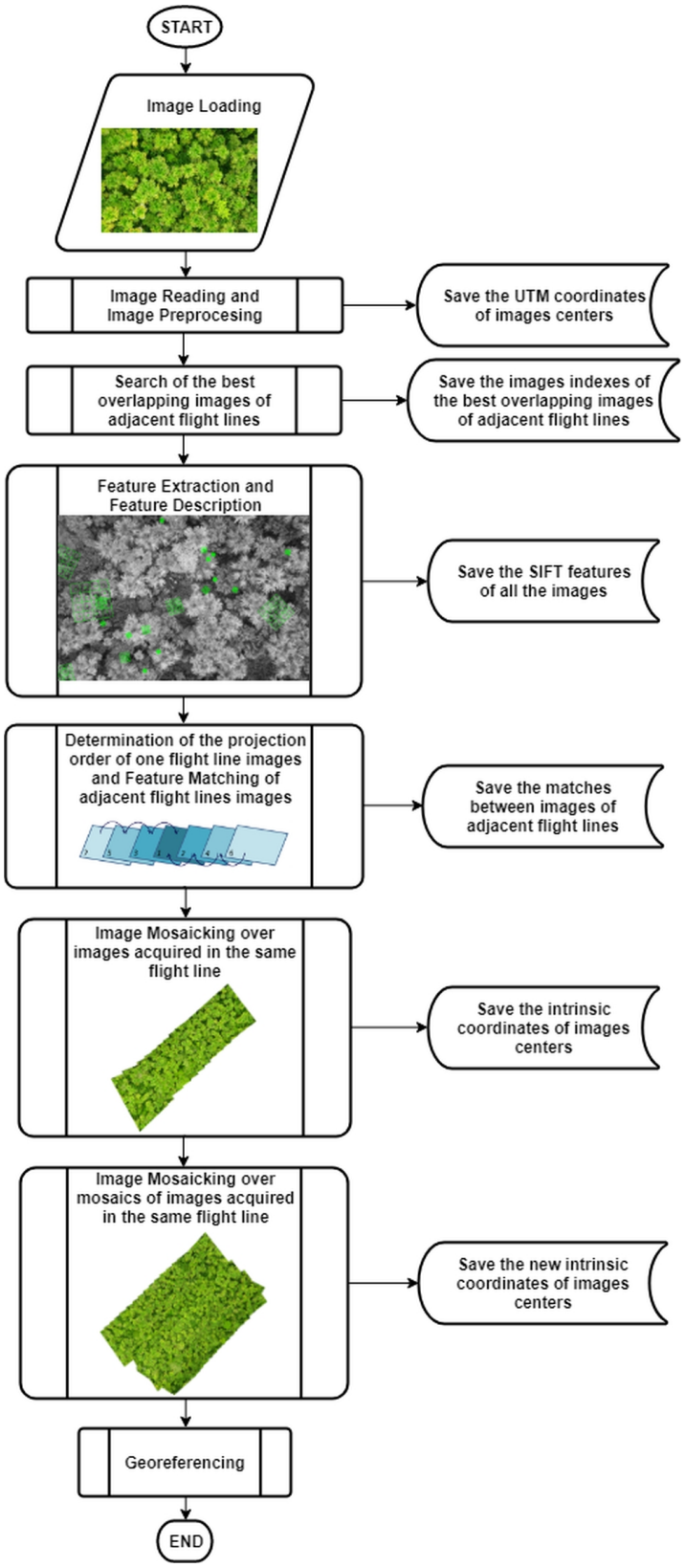


Figure 1: Flow Diagram of the proposed mosaicking method.

1. All the shooting distances *dist* are used to generate a histogram with equally spaced intervals.
2. The most common shooting distance $dist_{mr}$ is calculated as the mean value of the interval with the larger number of observations or greater frequency.
3. We define the parameter w as the ninth part of the difference between the largest and shortest shooting distance, according to (1). After evaluating the shooting distance for all the geotagged images of different flights, we found that dividing the difference between the boundaries in nine parts performs a better selection of images pertaining to the same flight line.

$$w = \frac{max(dist) - min(dist)}{9} \tag{1}$$

4. The range of distances is fixed by (2). The minimum value of this range is the most common shooting distance $dist_{mr}$ minus two times the parameter w, while the maximum value is the most common shooting distance $dist_{mr}$ plus five times the parameter w. This is mainly because the majority of shooting distances that do not correspond to images of the same flight line, indicated in the histogram of Figure 3 as red bars, are smaller than $dist_{mr}$. The choices of the multiplicative factors 2 and 5 were given after evaluating all the available geotagged images of different flights.

$$Range_{dist} = [dist_{mr} - 2w; dist_{mr} + 5w] \tag{2}$$

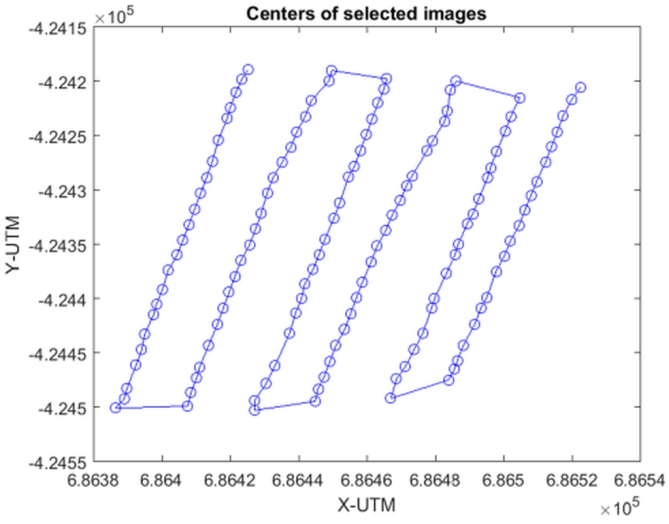


Figure 2: Plot of the geolocation of image centers selected for mosaicking.

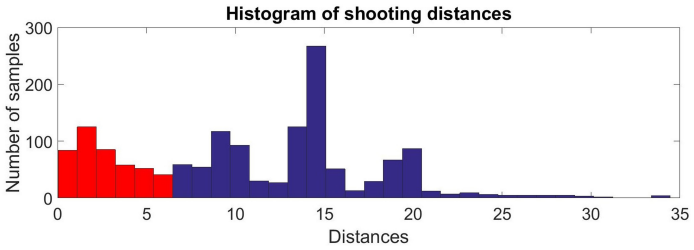


Figure 3: Histogram of shooting distances corresponding to six flights.

2.2 Search of the best overlapping images of adjacent flight lines

In order to perform the mosaicking of individual flight lines mosaics, it is required to compute image matching between images of adjacent lines. The optimal procedure is to search for the best overlapping images of adjacent lines rather than perform image matching for all the acquired images due to the time and computational cost. We utilize Delaunay Triangulation [11] and the UTM geographical coordinates of each image to search for pairs of images with enough overlap that belongs to adjacent flight lines. We applied the Delaunay Triangulation algorithm of Matlab that is basically an incremental implementation.

The fundamental property of Delaunay Triangulation can be explain based on the circles circumscribed to a triangle for the 2D case. Given four points V_1, V_2, V_3 and V_4, Delaunay Triangulation is fulfilled when each circle circumscribed to the triangle drawn from three points contains no other point inside of it. The importance of this property is that the majority of the drawn triangles possess relatively large internal angles. Thus, those triangles are considered to be "well shaped" (more similar to an equilateral triangle than to an obtuse one) and for our application, the edges of those triangles connect points that represent two images with enough overlap because the larger edges are smaller than the larger edges of triangles in the non-Delaunay Triangulation.

By applying Delaunay Triangulation, the connections between images with enough overlap are obtained, and shown in Figure 4a; nevertheless, those connections between images of adjacent lines are required to be selected and the rest, to be filtered out. In addition, all the edges of triangles with any internal angle larger than 90 degrees should be deleted to ensure better overlap and therefore more accurate matching points or inliers. After this process, the connections between the images are depicted in Figure 4b.

2.3 Feature Extraction and Feature Description

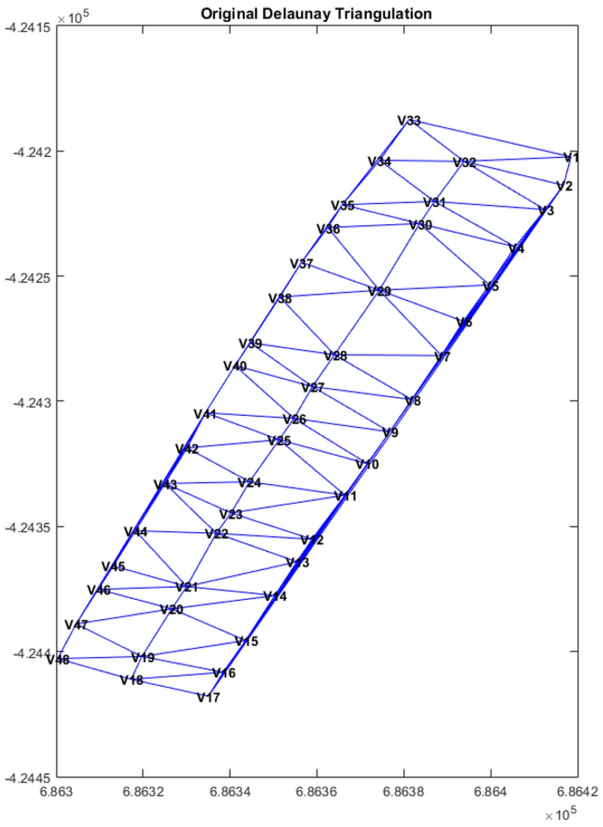
There are many feature extraction techniques in the literature but one of the best techniques used for mosaicking are the Scale Invariant Feature Transformation (SIFT) [12] because of the property to find corresponding points in different scale, rotated and translated images. Images in different scales are the result of convolution between an image $I(x, y)$ and 2D Gaussian filters $G(x, y, \sigma)$ with different values of standard deviations σ. Equation (3) describes how images in different scales $L(x, y, \sigma)$ are produced, while (4) describes the 2D Gaussian filter.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{3}$$

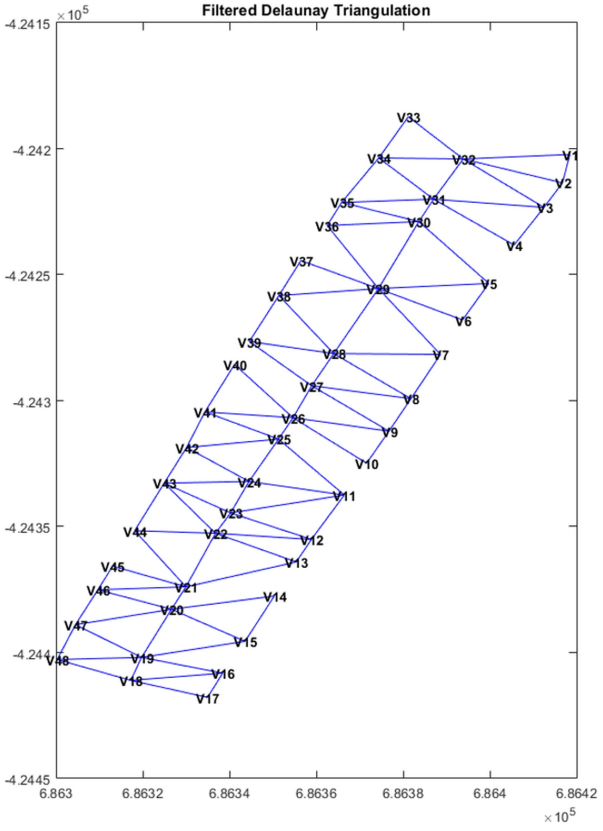
$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \tag{4}$$

SIFT features are located in the local maximum in a neighborhood of 3x3x3 (X-axis, Y-axis and the scale) over the result of the convolution between an image and the difference of the Gaussians with scales k and $k\sigma$, as expressed in (5). In addition, the gradients and its orientations are calculated in the neighborhood of the feature.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \tag{5}$$



(a)



(b)

Figure 4: Delaunay Triangulation used to search the best overlapping images of adjacent flight lines (a) Original Triangulation (b) Filtered Triangulation.

The first step is feature extraction, also known as feature detection. It computes the detection vector, which is composed of the location in the X-axis and Y-axis, the scale (represented by the standard deviation σ) and the resultant orientation of the gradients. Then, feature description computes the description vector, a 128-vector, from the gradients calculated in the neighborhood of the feature point.

Both feature extraction and description are computed by using the implementation of the SIFT algorithm in the open source library VLFeat [13]. Figure 5 shows some SIFT features computed over an aerial image acquired by an UAV over a swampy forest of the Peruvian Jungle in Iquitos, known as Aguajales [14-16].

2.4 *Determination of the projection order of one set of flight line images and Feature Matching of sets of adjacent flight lines images*

Performing image mosaicking over one set of flight line images requires certain implementation considerations because it implies the projection of images by the homographies, which estimation is not error free. Because of this, determining the projection order is a fundamental step that avoids generating mosaics with visual errors accumulated in certain direction of the image, as the one shown in Figure 6a.

The idea is to project all images over a projection plane aligned with the image located in the middle of the flight line. Therefore, the first image in the projection order is the image in the middle of the line. The next images in the projection order are those that are adjacent to the first one, and then it continues in the same manner but alternating between the two directions of the flight line. This criterion is depicted in Figure 6b.

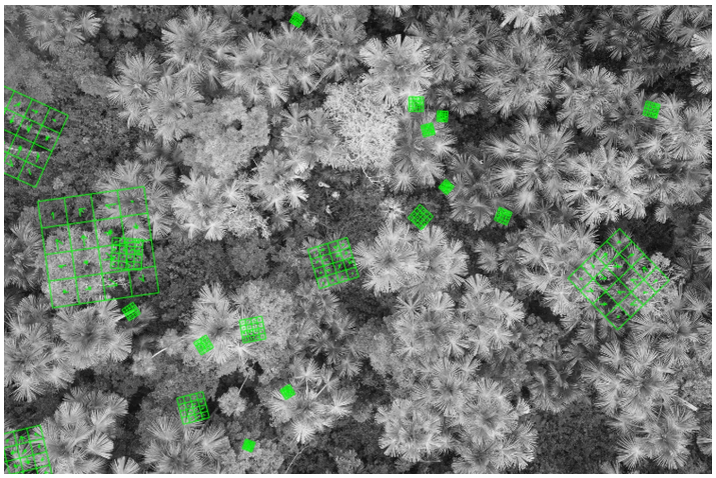
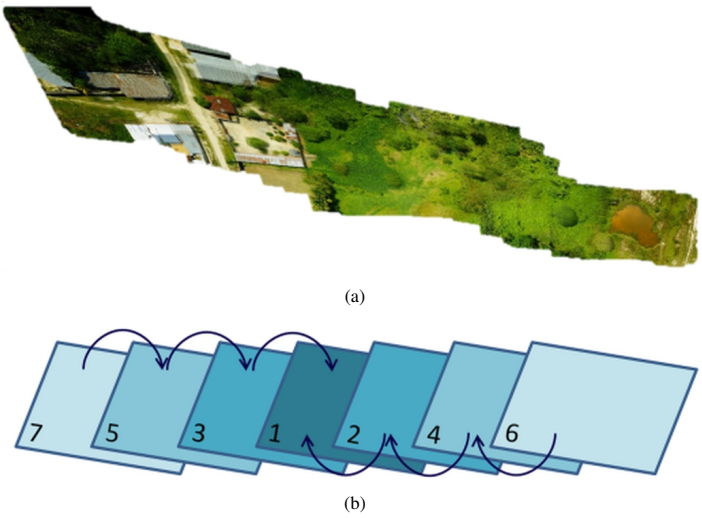


Figure 5: SIFT features computation over an aerial image acquired in a forest.

In this step of the proposed method, the feature matching is also performed, between images of adjacent flight lines that were selected in second step of this method. This operation finds two features, one per image, that possess similar SIFT descriptor vectors. We used the SIFT feature matching algorithm implemented in VLFeat and we obtained two results: the matches (pair of features) and the Euclidean distances between them.

The matches between two images are represented as a group of points X_c and X'_c; each match is represented as the homogeneous vectors x_c and x'_c. These homogeneous vectors are 3-vectors composed by the position of the feature in the X-axis and the Y-axis, and the unity. It is important to mention that the X-axis of the image refers to the horizontal axis (columns) and the Y-axis refers to the vertical axis (rows).

To find the best matches between images of adjacent flight lines, we select the 50 matches with the lowest Euclidean distance for every pair of image established in the second step of the proposed method. By doing this, more than 500 matches can be computed, between images corresponding to two adjacent flight lines that will be used to perform the calculation of homographies between mosaics of flight lines. Any transformation that will be applied to an image during the generation of mosaics of flight lines should be also applied to the features of each image corresponding to the matches between images of adjacent lines because they represent positions over images that are being modified.



(a)

(b)

Figure 6: Projection order of images: (a) Mosaic generated after projecting the images respect to an image in the end or beginning of line (b) Projection order proposed for image mosaicking over images in a flight line.

After several homography estimation tests, we observed that 50 matches should be computed to reduce the processing time and, at the same time, to ensure that the number of wrong matches or outliers is very small in comparison with the inliers; thus, the probability p_o of choosing a sample free of outlier increases.

2.5 *Image Mosaicking over images acquired in the same flight line*

This is the most important step of the proposed method. Both the homography estimation and the projection of images are performed following the projection order computed in the previous step (step 4th). To perform homography estimation, the feature matching between images of the same flight line is required. In this case, we select the 150 matches with lowest Euclidean distance instead of 50 because these matches are used to estimate the homography between individual images rather than mosaic images (mosaic of images acquired in the same flight line).

The homography H is a $3x3$ matrix that projects a set of points X_c belonging to the plane $P \in R^2$, as another set of points X'_c, belonging to another plane $P' \in R^2$. The projection of a point represented by a homogeneous vector x_c through the homography H is expressed in (6).

$$x'_c = Hx_c \tag{6}$$

One of the most common methods to perform homography estimation is based on the Random Sample Consensus or RANSAC method, and a posterior optimization which consists in minimizing a cost function depending on the homography and the inliers. This method is denoted as RANSAC-OPT in this work, for simplification purposes. RANSAC performs a robust estimation based on the calculations for a random sample by iteratively approximating a model. In each iteration, RANSAC computes not only the estimated model; in addition, it classifies all the matches. If a match adjusts to the estimated model with an error less than a threshold t, then it will be considered an inlier; otherwise, it will be considered an outlier.

We implemented RANSAC-OPT method for homography estimation and we observed that it gives good results with a tolerable error for images with an overlap of 60 percent or greater. Nevertheless, due to factors external to the flight planning, some images could be acquired with less overlap. In this situation, RANSAC-OPT can wrongly estimate the homography, producing deformations in mosaics, as those shown in Figure 7. Besides, this estimation can take more iterations to finish, thus increasing the processing time.

Because of these limitations of the RANSAC-OPT method, the proposed method verifies that the symmetric transfer error per inlier would be low, ensuring an acceptable homography estimation. The proposed algorithm of automatic homography estimation, depicted in Figure 8, is based on two modifications of the RANSAC algorithm followed by the optimization step. These two modifications are denoted RANSAC-1-OPT and RANSAC-2-OPT. The inputs of both modifications are the S matches between two images X_c y X'_c; the outputs are the estimated homography H_e, the percentage of inliers p_{inl}, and the symmetric transfer error per inlier ste_{inl}. The steps of this method are indicated below.



Figure 7: Deformation during mosaicking produced by the wrong estimation through RANSAC for two images with low overlap.

1. RANSAC-1-OPT is performed. This is composed of the first modification of RANSAC (RANSAC-1) and the posterior optimization (OPT).

(a) RANSAC-1: The iteration counter cnt_1 is initialized with a value of zero, and the number of required iterations N with a value different from zero. The probability p_o of choosing a sample free of outlier is also initialized with a value of 0.99 as described in the original RANSAC method.

i. $s = 4$ matched points are randomly selected, from the sample space of S matches, so that these s matches are as distant as possible. These s matches are used to perform the initial estimation of the homography through the Direct Linear Transformation (DLT) algorithm. This allows that the estimated homography can project appropriately any set of points dispersed in the whole image, rather than just points located in one portion of the image. In order to find the most distant matched points, the matches were randomly paired and then the two most distant pairs of matched points where chosen. The fact that $s = 4$ is because it is the number of matches needed to estimate the homography H_o by using the DLT method.

ii. Given the S homogeneous vectors x_c and x'_c (representing the matches) and the estimated homography H_o, the solution of the minimization problem of (7) is a homography H_R that minimizes the symmetric transfer error for all the S matches. The cost function of this minimization problem is the sum of the symmetric transfer error per match ste_c. Equation (8) is used to calculate ste_c, where $d(.)$ denotes the Euclidean distance between two points.

$$H_R = \min_{H_o} \sum_{i=1}^{S} ste_{c_i} \tag{7}$$

$$ste_{c_i} = d(x_{c_i}, H_o^{-1} x'_{c_i})^2 + d(x'_{c_i}, H_o x_{c_i})^2 \tag{8}$$

iii. Using the homography H_R, the errors of each match with respect to the estimated model (the homography) are represented by the values of ste_c, so they are calculated by using (8). Then the standard deviation of those errors σ_c are computed.

iv. The inliers of the model are represented by the homogeneous vectors x_{inl} and x'_{inl}, and they are those matches whose values of ste_c are less than the threshold t, computed according to (9). Once the inliers are identified, it is also computed the number of inliers n_{inl}, the sum of symmetric transfer errors of the inliers STE_{inl}, and the standard deviation of the symmetric transfer errors of the inliers σ_{inl}.

$$t = \sqrt{5.99}\sigma_c \tag{9}$$

v. During the first iteration, the estimated homography H_R, the homogeneous vector of the inliers x_{inl} and x'_{inl}, the number of inliers n_{inl}, the sum of symmetric transfer errors of the inliers STE_{inl}, and the standard deviation σ_{inl}, are save as reference for the following iterations. For the rest of iterations, the following instructions are performed:

- If the calculated STE_{inl} is less than or equal to 10 times its reference, the reference will be updated with the calculated STE_{inl}, and then it will proceed with the next item.
- If the computed n_{inl} is greater than its reference, the reference for the values of H_R, n_{inl}, STE_{inl}, σ_{inl}, x_{inl} and x'_{inl} will be updated with the computed values. Otherwise, it will proceed with the next item.
- If the computed n_{inl} is equal to its reference, the standard deviation σ_{inl} computed will be compared with the respective reference. If the calculated σ_{inl} is less than or equal to its reference, the reference values of H_R, n_{inl}, STE_{inl}, σ_{inl}, x_{inl} and x'_{inl} will be updated with the computed values. Otherwise, it will proceed with the step 1-(a)-vi.

vi. The iteration finishes and the iteration counter cnt_1 increases in 1. Then the number of required iterations N is computed adaptively based on the number of inliers n_{inl}. Equation (10) is used to make this calculation. The probability of choosing an outlier ϵ is calculated using (11).

$$N = \frac{\log(1 - p_o)}{\log(1 - (1 - \epsilon^s))} \tag{10}$$

$$\epsilon = 1 - \frac{n_{inl}}{S} \tag{11}$$

vii. When the iteration counter cnt_1 would be greater than or equal to N, RANSAC-1 finishes and its outputs are the reference values for H_R, n_{inl}, STE_{inl}, σ_{inl}, x_{inl} and x'_{inl}.

(b) OPT: The homography obtained after RANSAC-1 H_R is re-estimated utilizing just the n_{inl} homogeneous vectors representing the inliers x_{inl} and x'_{inl}. The output is the homography H_e, which solve the minimization problem of (12). The cost function is the sum of the symmetric transfer error of all the inliers, presented in (13).

$$H_e = \min_{H_R} \sum_{i=1}^{n_{inl}} ste_i \tag{12}$$

$$ste_i = d(x_{inl_i}, H_R^{-1} x'_{inl_i})^2 + d(x'_{inl_i}, H_R x_{inl_i})^2 \tag{13}$$

(c) Given the estimated homography H_e and the inliers, (14) is used to calculate the percentage of inliers p_{inl}, and (15) is employed to compute the symmetric transfer error per inlier ste_{inl}.

$$p_{inl} = \frac{n_{inl}}{S} \tag{14}$$

$$ste_{inl} = \frac{\sum_{i=1}^{n_{inl}} ste_i}{n_{inl}} \tag{15}$$

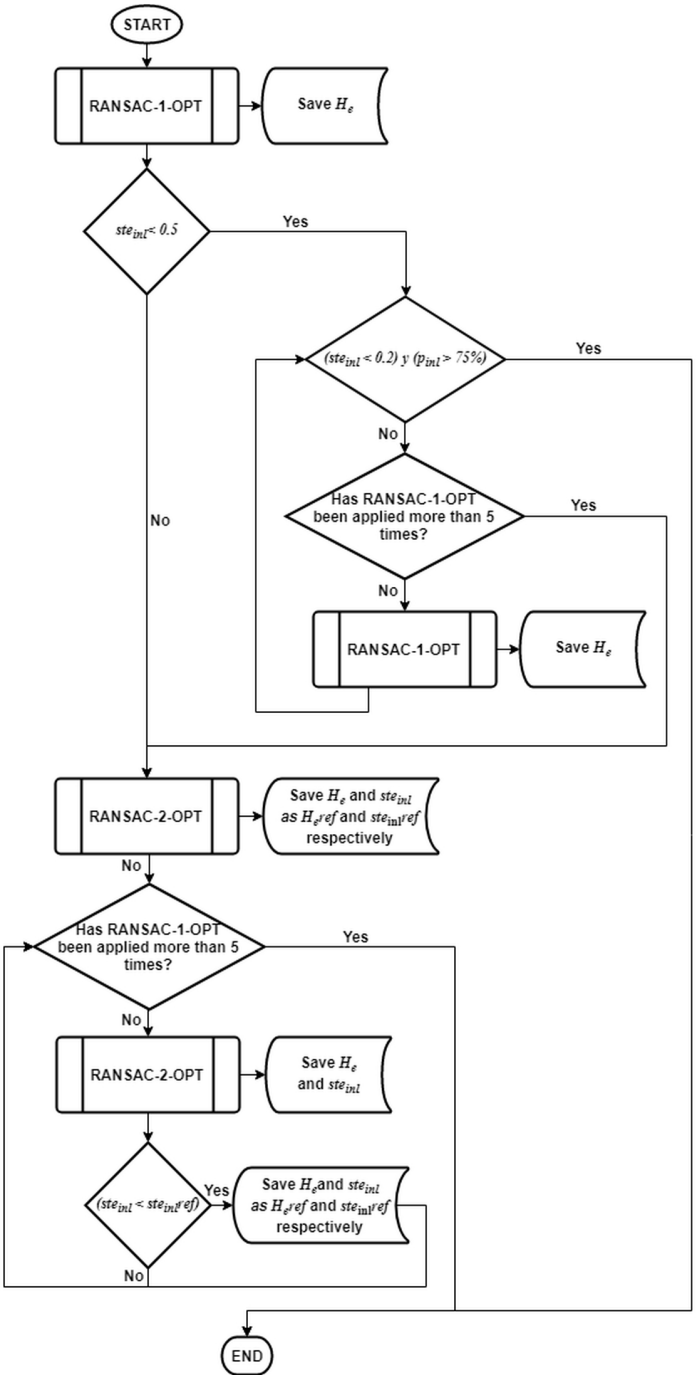


Figure 8: Proposed algorithm for the automatic estimation of homographies.

2. If the ste_{inl} obtained in the step 1 is less than 0.5 pixels, it will proceed with the following item. It is important to mention that the value of 0.5 pixels was established after several tests of homography estimation between images with low overlap.

   (a) If the ste_{inl} is less than 0.2 and the p_{inl} obtained in the step 1 is greater than 75%, the estimated homography H_e will be the solution and the proposed method for automatic estimation of homography will end. Otherwise, it will proceed with the next item. The threshold value for the percentage of inliers p_{inl} was also established empirically.

   (b) RANSAC-1-OPT (described in the step 1) is applied again until the conditions of the previous item are satisfied, or until RANSAC-1-OPT has been applied 5 times in total. We applied continuously RANSAC-1-OPT several times and we found that the result of the homography estimation did not improve after 5 times, so we established this threshold. If the solution is not found after applying RANSAC-1-OPT 5 times, it will proceed with the step 3.

3. RANSAC-2-OPT is performed. This is composed of the second modification of RANSAC (RANSAC-2) and the posterior optimization (OPT).

   (a) RANSAC-2: It initializes the same parameters of RANSAC-1 with the same values: the iteration counter cnt_2, the number of required iterations N, and the probability p_o of choosing a sample free of outlier.

      i. $s = 4$ matched points are randomly selected, from the sample space of S matches, to perform the initial estimation of the homography through the DLT algorithm. In this case, the matches are randomly selected without any other condition.

      ii. Given the S homogeneous vectors x_c and x'_c and the estimated homography H_o, the symmetric transfer error per match ste_c is computed using (8). The standard deviation of those errors σ_c is also computed.

      iii. In the same manner as for RANSAC-1, the homogenous vectors of inliers x_{inl} and x'_{inl}, the number of inliers n_{inl}, and the standard deviation of the symmetric transfer error of the inliers σ_{inl} are calculated.

      iv. During the first iteration, the estimated homography H_o, the homogeneous vector of the inliers x_{inl} and x'_{inl}, the number of inliers n_{inl}, and the standard deviation σ_{inl}, are saved as reference for the following iterations. For the rest of iterations, the following instructions are performed:

         - If the computed n_{inl} is greater than the reference value of n_{inl}, the reference values of H_o, n_{inl}, σ_{inl}, x_{inl} and x'_{inl} will be updated. Otherwise, it will proceed with the next item.
         - If the computed n_{inl} is equal to its reference, the standard deviation σ_{inl} computed will be compared with the respective reference. If the computed σ_{inl} is less than or equal to its reference, the reference values of H_o, n_{inl}, σ_{inl}, x_{inl} and x'_{inl} will be updated with the computed values. Otherwise, it will proceed with the step 3-(a)-v.

      v. The iteration finishes and the iteration counter cnt_2 increases in 1. Then the number of required iterations N is computed adaptively as it was done for RANSAC-1.

vi. When the iteration counter cnt_2 would be greater than or equal to N, RANSAC-2 finishes and its outputs are the reference values for H_o, n_{inl}, σ_{inl}, x_{inl} and x'_{inl}.

(b) OPT: The homography obtained after RANSAC-2 H_o is re-estimated utilizing just the n_{inl} homogeneous vectors representing the inliers x_{inl} and x'_{inl}. This is performed in the same way as performed in the step 1-(b), and the result is the homography H_e.

(c) Given the estimated homography H_e and the inliers, the percentage of inliers p_{inl} and the symmetric transfer error per inlier ste_{inl} are computed as it was done in the step 1-(c).

4. The homography H_e and the symmetric transfer error per inlier ste_{inl} obtained in the step 3 are saved as references H_e^{ref} and ste_{inl}^{ref}. Then, RANSAC-2-OPT is applied again and the calculated ste_{inl} is compared with the reference ste_{inl}^{ref}. If the calculated value is less than the reference, then the references H_e^{ref} and ste_{inl}^{ref} will be updated. This procedure is repeated 4 times.

The H_e^{ref} obtained after the repetitions is the solution and the proposed algorithm for the automatic estimation of the homography finishes.

An important detail of the implementation of RANSAC-1 and RANSAC-2 is the normalization of the homogeneous vectors of the matches. Before using these vectors for the homography estimation, it is required to normalize them; at the same time, denormalization must be performed after the homography estimation in order to maintain the original scale. Both normalization and denormalization improve the results by avoiding high errors during homography estimation.

The algorithm to calculate adaptively the number of iterations, the DLT algorithm, the normalization algorithm, and the original RANSAC algorithm are well developed in [17].

Once the homographies between two adjacent images in the same flight line are computed, the homographies between any image of the line and the central image (first image in the projection order) can be computed by multiplying homographies according to the sequence established and depicted in Figure 6b.

The homographies allow the computation of the location of images over the mosaic images, and then, the intensities of pixels are interpolated for the new positions. Merging all the projected images implies to find the intensity of pixel over the mosaic by averaging (or applying another similar procedure) the intensity of the same pixel in the projected images.

The new position of the images centers, and the position of the matches points between images of adjacent flight lines should be computed for each projected images through the homographies.

2.6 Image Mosaicking over mosaics of images acquired in the same flight line

This step of the proposed mosaicking method performs the image mosaicking of the mosaics obtained from images acquired in the same flight lines. The homographies between those individual mosaics are computed through the matches from images of adjacent lines. It proceeds as follows:

1. From the new positions of matches from images of adjacent lines, the matches from two adjacent mosaics X_m and X'_m are computed.

2. The sequence of projection is determined in the same manner that was done for the images belonging to same flight lines (step 4th) so the first image in the sequence is the mosaic image located in the middle.

3. The proposed algorithm for automatic estimation of homography (step 5th) is applied to compute the homography between mosaic images.

4. The positions of the mosaic images in the whole mosaic is calculated, and the pixels information is interpolated for the new positions. This procedure is the same performed for individual images in the previous step (step 5th).

5. The new positions of the images centers are computed by using the estimated homographies. Georeferencing requires the calculation of the new position of the images centers.

2.7 Georeferencing

In order to proceed with the georeferencing of the final mosaic, the position of the images centers as well as the UTM coordinates are required. The following steps detail this procedure:

1. The orientation of the Y-axis of the image is inverted to make possible align the image axes with the UTM coordinates axes through a rotation. Given the number of rows in the image M and the coordinates in Y-axis y, the new coordinates in the inverted Y-axis y' are calculated by (16).

$$y' = M + 1 - y \tag{16}$$

2. The angle between the image axes and the UTM coordinates axes α_{rot}, shown in Figure 9, is computed. Given the intrinsic coordinates (image coordinates) of the images centers belonging to the first two images in the projection order (Step 4th) of the central line (xc_1, yc'_1), and (xc_2, yc'_2), and the UTM coordinates of the same images ($xutm_1$, $yutm_1$) and ($xutm_2$, $yutm_2$); the angle between the line that connect the centers of these two images and the X-axis of the image α_{pix} is calculated using (17). Besides, the angle between the line that connect the centers of these two images and the X-axis of UTM coordinate system α_{UTM} is computed using (18). The angle α_{rot} is computed using (19).

$$\alpha_{pix} = \arctan\left(\frac{yc'_2 - yc'_1}{xc_2 - xc_1}\right) \tag{17}$$

$$\alpha_{UTM} = \arctan\left(\frac{yutm_2 - yutm'_1}{xutm_2 - xutm_1}\right) \tag{18}$$

$$\alpha_{rot} = \alpha_{pix} - \alpha_{UTM} \tag{19}$$

3. The image of the final mosaic is rotated $-\alpha_{rot}$ in the clockwise direction to align the axes of the image and the axes of the UTM coordinates.

4. After rotating the final mosaic image, the intrinsic coordinates of images centers will vary. Because of this, it is required to calculate these new rotated coordinates from the distances between centers of individual images and the center of the final mosaic, the rotated coordinates of the center of the final mosaic, and the rotated angle.

5. Linear regression is performed from the rotated coordinates of images centers and the UTM coordinates of the centers. Given the coordinates of images centers (xc, yc'), the UTM coordinates of the centers ($xutm$, $yutm$) can be expressed based on the coefficients of the linear regression a_1, a_2, a_3, a_4, a_5 and a_6 according to (20) and (21).

$$xutm = a_1 xc + a_2 yc' + a_3 \tag{20}$$

$$yutm = a_4 xc + a_5 yc' + a_6 \tag{21}$$

6. The UTM coordinates of the four corners of the image are computed from (20), (21) and the coefficients a_1, a_2, a_3, a_4, a_5 and a_6.

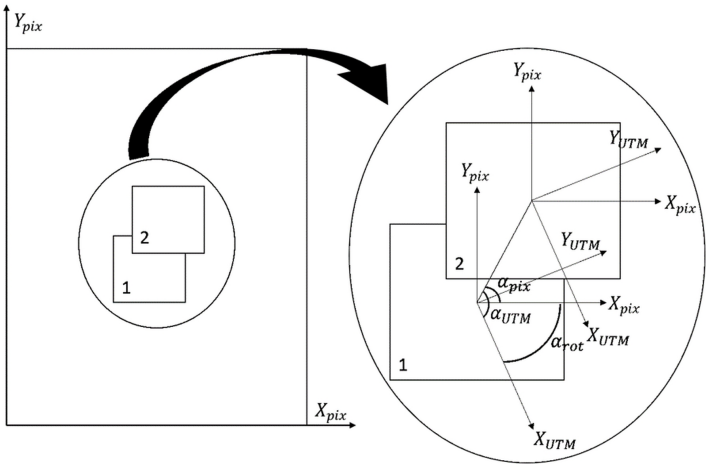


Figure 9: Angle between the image axes and the axes of the UTM coordinates.

3 Results and Discussion

We describe the results of applying the proposed method over 6 sets of images acquired in 6 different flights regarding two scenarios: a city and a forest in Peruvian Amazon. These images were acquired with the professional camera Sony NEX-7, which allows taking images of 24.3MP, at a height of 100 meters from the ground. Each image possesses a spatial resolution of 6000x4000 pixels; but their sizes are reduced to a quarter of the original size during the processing, as it was explained in the first step of the proposed method.

The proposed method was implemented in MATLAB R2016a, and for validation purposes, we utilized the professional software Pix4DMapper Pro version 3.2.10, whose algorithm is based on Structure from Motion to generate 3D points and then orthorectified the mosaic as explained in [18]. We compared the processing time and the geolocation errors obtained in the same computer: HP Z820 Workstation with two processors Intel Xeon E5-2690 2.9GHz and 96 GB RAM. It is important to mention that the software Pix4DMapper Pro also processed the images resized to a quarter of the original size.

The processing time required to obtained each mosaic image is shown in Table 1. Pix4DMapper Pro performs a preprocessing step to calibrate images and it uses just the calibrated images for the image mosaicking; therefore, the processing time for this process is not included in our analysis.

Table 1: Processing Time

Flight	Number of images	Pix4DMapper Pro	Proposed Method
1	73	10 min 29 s	10 min 36 s
2	145	24 min 44 s	21 min 37.2 s
3	148	24 min 20 s	24 min 15 s
4	168	30 min 06 s	23 min 18 s
5	168	28 min 46 s	27 min 13.5 s
6	138	20 min 46 s	27 min 46 s

In addition, the geolocation errors corresponding to the UTM X-axis and the UTM Y-axis have been computed. From those errors, we calculate the mean absolute error (MAE) and the root mean square error (RMS) which are shown in Table 2.

Table 2: Geolocation Errors

Flight	Axes	Pix4DMapper Pro		Proposed Method	
		MAE (m)	RMS (m)	MAE (m)	RMS (m)
1	X	0.9341	1.5209	1.2280	1.5942
	Y	3.7711	4.8955	3.4296	4.0248
2	X	1.9233	2.3048	2.5927	3.1386
	Y	4.3136	5.4267	4.0290	4.5155
3	X	2.2737	2.6818	3.4259	3.6941
	Y	4.6558	5.7388	8.1738	9.1930
4	X	4.2449	4.9988	3.9255	4.4720
	Y	3.7049	4.3406	3.1868	3.7808
5	X	3.6335	4.3212	4.4445	4.9658
	Y	3.8871	4.6748	2.8436	3.2852
6	X	0.8314	1.3360	1.7022	2.0033
	Y	4.6199	5.6286	3.7153	4.1984

From Table 1, we can observe that for flights 2, 3, 4 and 5 (corresponding to a forest in Peruvian Amazon), the proposed method was 6 minutes and 48 seconds faster in the best case, and just 5 seconds faster in the worst case. On the other side, for the flights 1 and 6, Pix4DMapper Pro was 7 minutes faster in the best case, and just 7 seconds faster in the worst case.

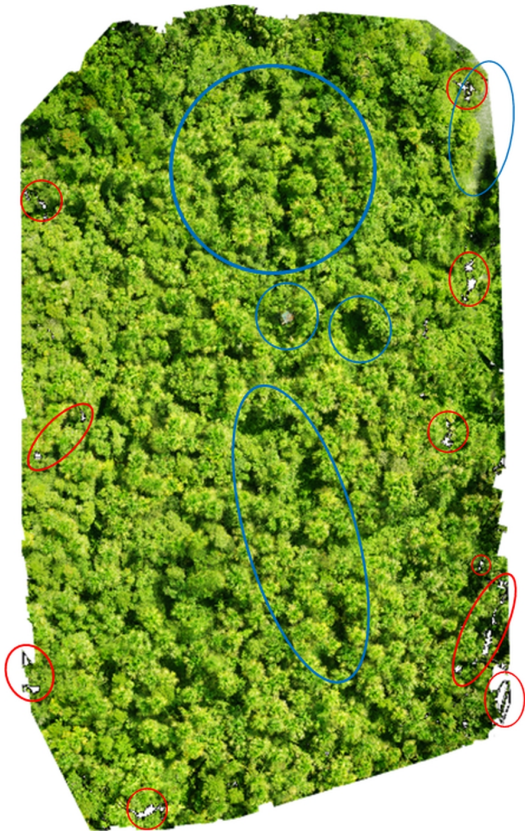


(a)

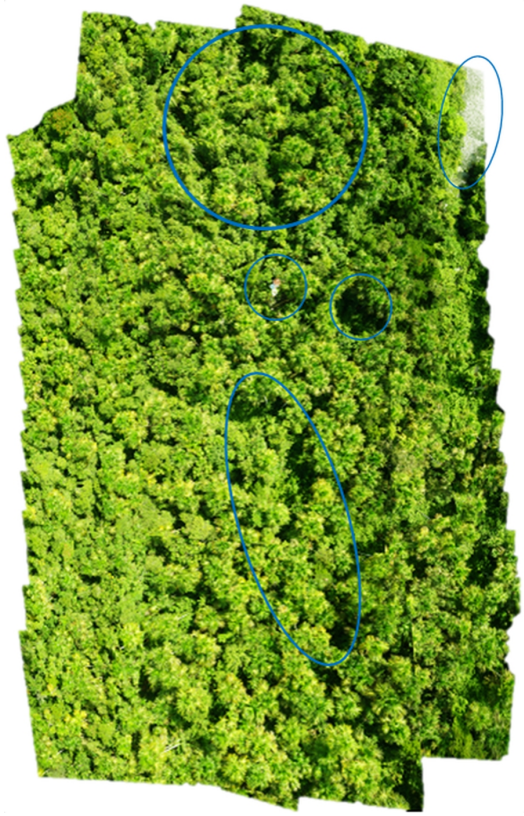


(b)

Figure 10: Image mosaics obtained for the flight 1: (a) Pix4DMapper Pro (b) Proposed Method.



(a)



(b)

Figure 11: Image mosaics obtained for the flight 6: (a) Pix4DMapper Pro (b) Proposed Method.

Regarding the MAE and RMS geolocation errors of the flights 1, 2, 5 and 6, the geolocation errors (both MAE and RMS) in X-axis computed with the proposed method were greater than those obtained by Pix4DMapper Pro; while, the geolocation errors in Y-axis computed with the proposed method were less than those obtained by Pix4DMapper Pro.

In the case of flight 3, MAE and RMS in both axes computed with our method were greater than those computed with Pix4DMapper Pro; but in flight 4, MAE and RMS in both axes computed with our method were less than those computed with Pix4DMapper Pro. For the majority of flights, the difference of both geolocation errors are low.

We observe that the geolocation errors computed with both methods in the Y-axis are always greater than the computed in the X-axis because of the employed GPS has that horizontal error distribution for all its measurements.

The highest differences in processing time and geolocation errors observed in Table 1 and Table 2 are because the inliers and the estimated homography for the same pair of images may produce different errors and different processing time. These differences are typical of the homography estimation step.

In Figure 10 and Figure 11, we compared the mosaic images obtained with Pix4DMapper Pro and our method for the flight 1 and 6 respectively. In the mosaic image of flight 1, regarding to a city, we can observe more clearly the differences between image mosaics. The image mosaic obtained by Pix4DMapper Pro presents some small holes in the boundaries of the image (indicated with red circles and ellipses). Besides, some trimmed edges and other differences caused by the enhancement of interpolated pixels information performed by Pix4DMapper Pro are encircled by blue ellipses. The mosaic image obtained with our method does not present holes and trimmed edges. For flight 6, the differences are not so clear as those of Figure 10 because it is difficult to see these differences in a scenario like a forest. Visible differences in Figure 11 are shown in the same way that was done for Figure 10.

4 Conclusions

The proposed method requires images acquired when the camera points the nadir, so an stabilization system or gimbal should be used. In this way, a mosaicking algorithm without any preprocessing (like the one performed in the algorithm of Pix4DMapper Pro) can be utilized to obtain good results. Otherwise, a more complicated algorithm should be used to perform the estimation of position and orientation of the cameras.

Finally, our mosaicking algorithm based on a simple process results being viable and reproducible to perform image mosaicking of aerial images for applications of natural resources monitoring because it obtains similar geolocation errors and processing time in comparison with a professional software based in Structure from Motion; moreover, our results do not present any holes or color incoherencies.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment We thank the National Innovation Program for Competitiveness and Productivity, Innóvate Perú, for the funds allocated to our project (393-PNICP-PIAP-2014).

References

[1] A. Chaves, P. Cugnasca, J. Jose "Adaptive search control applied to Search and Rescue operations using Unmanned Aerial Vehicles (UAVs)" IEEE Latin America Transactions, **12**(7), 1278–1283, 2014. https://doi.org/10.1109/TLA.2014.6948863

[2] J. Everaerts, "The use of Unmanned Aerial Vehicle (UAVS) for remote sensing and mapping" in 2008 XXI ISPRS Congress Technical Commission I: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII Part B1, Beijing, China, 2008.

[3] Z. Li, V. Isler, "Large Scale Image Mosaic Construction for Agricultural Applications" IEEE Robotics and Automation Letters, **1**(1), 295–302, 2016. https://doi.org/10.1109/LRA.2016.2519946

[4] A. Iketani, T. Sato, S. Ikeda, M. Kanbara, N. Nakajima, N. Yokoya "Video mosaicing based on structure from motion for distortion-free document digitization" in 2007 Computer Vision ACCV 2007. Lecture Notes in Computer Science. Vol. 4844, Tokyo, Japan, 2007. https://doi.org/10.1007/978-3-540-76390-1_8

[5] Y. Yang, G. Sun, D. Zhao, B. Peng "A Real Time Mosaic Method for Remote Sensing Video Images from UAV" Journal of Signal and Information Processing, **4**(3), 168–172, 2013. https://doi.org/10.4236/JSIP.2013.43B030

[6] T. Botterill, S. Mills, R. Green "Real-time aerial image mosaicing" in 2010 25th International Conference of Image and Vision Computing New Zealand, Queenstown, New Zealand, 2010. https://doi.org/10.1009/IVCNZ.2016.6148850

[7] E. Garcia-Fidalgo, A. Ortiz, F. Bonnin-Pascual, J. Company "Fast image mosaicing using incremental bags of binary words" in 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 2016. https://doi.org/10.1109/ICRA.2016.7487247

[8] D. Brito, C. Nunes, F. Padua, A. Lacerda "Evaluation of Interest Point Matching Methods for Projective Reconstruction of 3D Scenes" IEEE Latin America Transactions, **14**(3), 1393–1400, 2016. https://doi.org/10.1109/TLA.2016.7459626

[9] T. Kekec, A. Yildirim, M. Unel "A new approach to real-time mosaicing of aerial images" Robotics and Autonomous Systems, **62**(12), 1755–1767, 2014. https://doi.org/10.1016/j.robot.2014.07.010

[10] Y. Xu, J. Ou, H. He, X. Zhang, J. Mills "Mosaicing of Unmanned Aerial Vehicle Imagery in the Absence of Camera Poses" Remote Sensing, **8**(3), 204–219, 2016. https://doi.org/10.3390/RS8030204

[11] A. Moussa, N. El-Sheimy, "A fast approach for stitching of aerial images" in 2016 XXIII ISPRS Congress, Commission III: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XLI Part B3, Prague, Czech Republic, 2016. https://doi.org/10.5194/isprsarchives-XLI-B3-769-2016

[12] D. Lowe "Disctintive image features from scale-invariant keypoints" International Journal of Computer Vision, **60**(2), 91–110, 2004. https://doi.org/10.1023/B:VISI.0000029664.99615.94

[13] A. Vedaldi, B. Fulkerson "VLFeat: An open and portable library of computer vision algorithms" in Proc. 18th International Conference on Multimedia, Firenze, Italy, 2010. https://doi.org/10.1145/1873951.1874249

[14] D. Del Castillo, E. Otárola, L. Freitas, Aguaje: the amazing palm tree of the Amazon, IIAP, Iquitos, Peru, 2006.

[15] L. Freitas, M. Pinedo, C. Linares, D. Del Castillo, Descriptores para el Aguaje (Mauritia flexuosa L. f.), IIAP, Iquitos, Peru, 2006.

[16] J. Levistre, J. Ruiz "El Aguajal: El bosque de la vida en la Amazonía Peruana" Ciencia Amazónica (Iquitos), **1**(1), 31–40, 2011. https://doi.org/10.22386/ca.v1i1.6

[17] R. Hartley, A. Zisserman, Multiple view geometry in computer vision, 2th edition, Cambridge University Press, 2003.

[18] O. Küng, C. Strecha, A. Beyeler, J. Zufferey, D. Floreano, P. Fua, F. Gervaix, "The accuracy of automatic photogrammetric techniques on ultra-light UAV imagery" in International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g): The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVIII Part C22, Zurich, Switzerland, 2011. https://doi.org/10.5194/isprsarchives-XXXVIII-1-C22-125-2011

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 2, 328-337 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Extending Movable Surfaces with Touch Interaction Using the VirtualTablet: An Extended View

Adrian H. Hoppe[*,1], Felix Marek[1], Florian van de Camp[2], Rainer Stiefelhagen[1]

[1]*Karlsruhe Institute of Technology (KIT), Institute for Anthropomatics and Robotics (IAR), cv:hci Lab, 76131 Karlsruhe, Germany*

[2]*Fraunhofer IOSB, Interactive Analysis and Diagnosis (IAD), 76131 Karlsruhe, Germany*

ARTICLE INFO

Article history:
Received: 07 January, 2020
Accepted: 04 February, 2020
Online: 30, March, 2020

Keywords:
VirtualTablet
Virtual environment
Virtual reality
Touch interaction
Haptic feedback

ABSTRACT

Immersive output and natural input are two core aspects of a virtual reality experience. Current systems are frequently operated by a controller or gesture-based approach. However, these techniques are either very accurate but require an effort to learn, or very natural but miss haptic feedback for optimal precision. We transfer ubiquitous touch interaction with haptic feedback into a virtual environment. To validate the performance of our implementation, we performed a user study with 28 participants. As the results show, the movable and cheap real world object supplies an accurate touch detection that is equal to a laserpointer-based interaction with a controller. Moreover, the virtual tablet can extend the functionality of a real world tablet. Additional information can be displayed in mid-air around the touchable area and the tablet can be turned over to interact with both sides. Therefore, touch interaction in virtual environments allows easy to learn and precise system interaction and can even augment the established touch metaphor with new paradigms.

1 Introduction

Virtual Reality (VR) has established itself at a consumer level. Many different VR systems immerse the user in an interactive environment. However, the different systems all have their distinct input devices. For newer users, we witnessed, it is quite difficult to just *pull the trigger* or *press the grip button*, since users do not see their own hands, but only the floating input device. More natural interfaces are needed. Gesture interaction via a camera that is mounted on the Head Mounted Display (HMD) lets the user grab virtual objects and manipulate the Virtual Environment (VE) effortlessly. But, haptic sensations are missing.
We therefore present VirtualTablet - A touchable object that is made of very cheap materials and can have any size or shape. This paper is an extension of work originally presented in the IEEE Conference on Virtual Reality and 3D User Interfaces (VR) [1]. It presents further detail regarding related work, implementation and design as well as evaluation and results. This paper makes two technical contributions: First, VirtualTablet supplies touch interaction on a movable surface captured by a movable camera. Second, different interaction techniques are presented that increase the functionality of the tablet to more than the capability of a real-world tablet.

2 Related Work

Haptic feedback increases performance and usability [2]–[4]. Different systems can be used to realize haptic feedback for a user. With active haptics, users wear a device that exerts pressure or a force on the user's hand. Scheggi et al. [5] use small tactile interfaces at the fingertips. The interfaces press down on the fingers if an object is touched. The VRGluv[1] is an exoskeleton that allows to constrain the movement of the fingers, thereby providing haptic feedback. With passive haptics the user touches real objects to receive feedback. The real objects can be tracked, for example with markers, and then displayed in the VE. Simeone et al. [6] track different real objects, e.g. a cup or an umbrella, as well as the head and hands of the user. They showed that the virtually displayed object does not need to match the real object perfectly. Some variation of material properties or shape allow the use of a broader range of virtual objects even if the set of real world objects is limited.

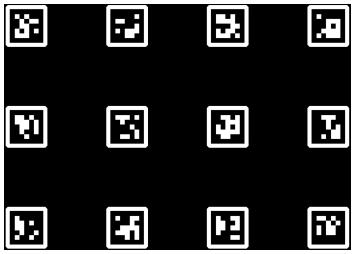
[*]Adrian H. Hoppe, Fraunhoferstr. 1, 76131 Karlsruhe, Germany, +49 721 6091 307 & adrian.hoppe@kit.edu
[1]see https://vrgluv.com/

If users start navigating in the VE, a mismatch between the location of the real and virtual objects may occur. To overcome this, movable haptic proxies or redirection techniques can be used. Araujo et al. [7] use a robotic arm to move a haptic surface to block the user's hand when it gets close to a virtual object. The head of the robotic arm can be rotated and exchanged, hereby providing a wide range of haptic surfaces. i.e. different textures, a pressure sensor, a interactive surface, physical controls or even a heat emitter. Kohli [8] implemented a touch interaction using an IMPULSE system that tracks the finger via a LED. Utilizing space warping, it is possible to alter the movement of the user's hand in the real world, without the user noticing. By steering the hand towards a real world object, haptic feedback can be given by a virtual object at another position. Azmandian et al. [9] track cubes and a wearable glove with markers on a table surface. The users hand and forearm, as well as some virtual cubes are displayed in the VE. By redirecting the hand of the user while reaching for the vritual objects, they provide haptic sensations for several virtual cubes with a single real world cube. Cheng et al. [10] apply the same principle. Using eye tracking and an OptiTrack system they predict what virtual object will be touched. The movement of the virtual hand is then modified so that a user corrects their real world movement. This allows Cheng el al. to only use a hemispherical surface as a touch proxy. Most of these techniques require a rather static scene and a non navigating user and are thereby not suitable for full body VR experiences.

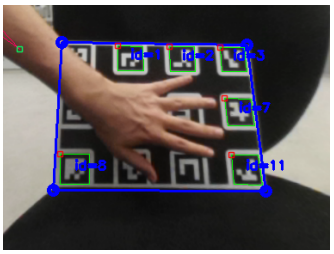
The presented systems above all give a haptic feeling to the user, but have limited input capabilities. To overcome this, Medeiros et al. [11] use a tracked tablet in a CAVE environment. Depending on the orientation of the tablet, the respective view into the virtual world is rendered on the touch screen. A user can select and manipulate objects by clicking on them in the rendered view or using a set of control buttons. The tablet is tracked by an external tracking system using markers. Xiao et al. [12] use the internal sensors of a Microsoft HoloLens to allow touch interaction on static surfaces. The system is much more flexible as all of the above systems, since it does not require the user to put on gloves or to set up an external tracking system. Using the depth sensors and a segmentation algorithm the hands of the user are recognized and the fingertips extracted. A RANSAC algorithm detects surfaces in the room. Combining these allows for touch interaction with haptic feedback on a real world surface. There is no need to visualize the hands or the surface for the user, since the HoloLens is an augmented reality (AR) HMD and the user sees the real objects. However, only static surfaces are used.

3 Implementation

The VirtualTablet system is implemented on a HTC Vive Pro VR HMD[2]. The Vive features a 1440×1600 pixels display per eye with a 90 Hz refresh rate. The field of view of the Vive is 110 °. Using the Lighthouse tracking system and an IMU the pose of the HMD is detected in a room area of $5 \times 5\,m^2$. Moreover, the Vive has a stereo camera on the front of the HMD[3]. Each camera has a resolution of 640×480 pixels with a refresh rate of 60 Hz and a field of view of 96° horizontal and 80° vertical. The camera can be accessed with the SRWorks toolkit with a latency of about 100 ms with SRWorks to 200 ms with the ZED Mini. SRWorks supplies depth information with an accuracy of +/- 3 cm at 1 m and +/- 10 cm at 2 m distance from the camera. The minimal distance is 30 cm.



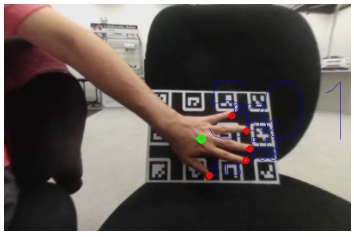
(a) ArUco board with 4×3 markers.



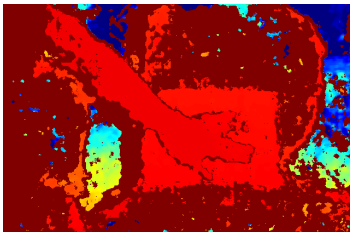
(b) Detected ArUco markers (green) and the derived board pose (blue).

Figure 1: Setup and Detection of the ArUco board.

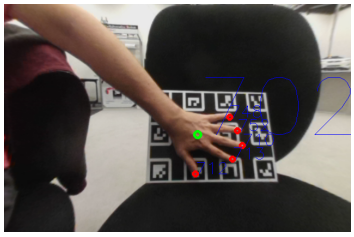
To detect the movable objects, ArUco markers [13] are used. They provide a position and rotation in respect to the camera, given their real world size. The VirtualTablet consists of a simple rigid base, for example a piece of acryl glass or cardboard, pasted up with ArUco markers. Because the hand of the user will overlap with some markers while interacting, the markers are spread out over the board (see Figure 1). ArUco markers can be uniquely identified, therefore allowing the use of more than one VirtualTablet at the same time.



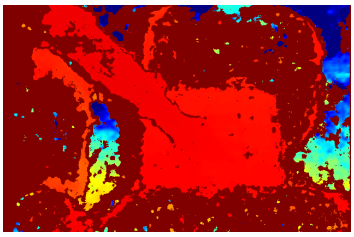
(a) Floating hand: left RGB camera.



(b) Floating hand: depth image.



(c) Resting hand: left RGB camera.



(d) Resting hand: depth image.

Figure 2: Front camera images of a hand hovering above the board and touching the board. Dark red color in the depth image mark invalid depth information.

To detect the fingertips, a segmentation algorithm is used, since the hand of the user cannot be reconstructed from the given depth information (see Figure 2). First, the hand is extracted from the RGB image by first masking out the board. An interacting hand will be inside this area. Since the ArUco board only consists of black and white colors, we subtract these colors from the image which results in a binary representation of the segmented hand. Pixels

[2] see https://www.vive.com/de/product/vive-pro/

[3] see https://developer.vive.com/resources/knowledgebase/intro-vive-srworks-sdk/ and https://www.stereolabs.com/blog/vive-pro-ar-zed-mini/

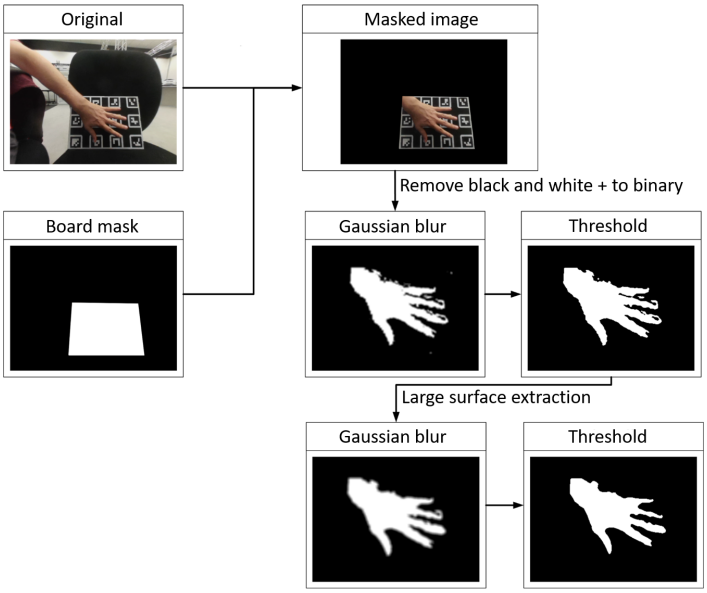


Figure 3: Hand segmentation process.

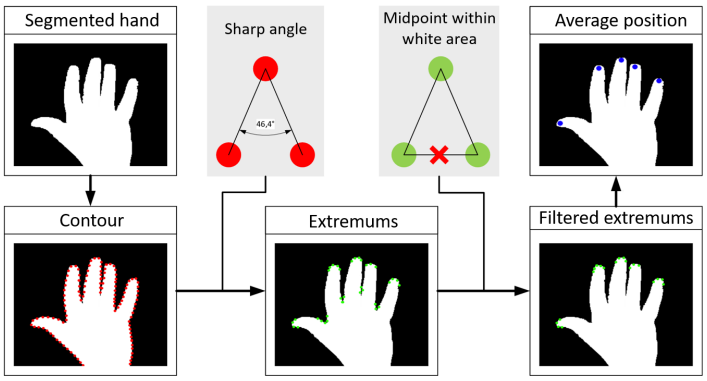
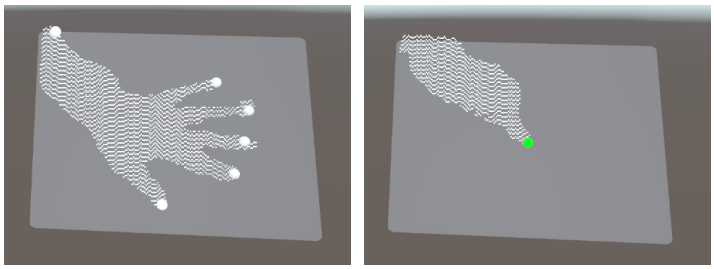


Figure 4: Fingertip detection process.

with a black or white color are removed in the HSV color space at a threshold with a low saturation (white) or a low brightness (black), i.e. $S \leq 80$ or $V \leq 40$ (scale from 0 to 255). As Figure 3 shows, the binary image is then enhanced by a blur filter, the application of a threshold and the extraction of larger surfaces[4]. There are other ways to extract the hand by using skin color, so this approach can be easily exchanged if the background colors or lighting conditions change. Second, the fingertips are detected by calculating the extreme values of the contour of the hand [14]. This is done by calculating the angle between three points on the contour that are for our camera setup 18 pixels apart from each other (see Figure 4). To track the fingertips from one frame to another a distance-based tracking algorithm similar to [12] is used. Combined with the depth information from the stereo camera the 2D tracked location of the user's fingers can be projected into the 3D space relative to the HMD. To have a more robust depth value an average of 11×11 pixels is calculated around the 2D fingertip location. Since the pose of the surface and the fingertips are known, a touch can be detected. If the finger is closer than 0.5 cm to the surface, a click is triggered.

[4]see https://medium.com/@muehler.v/simple-hand-gesture-recognition-using-opencv-and-javascript-eb3d6ced28a0



(a) Point cloud of the hand with detected fingertips (white spheres).

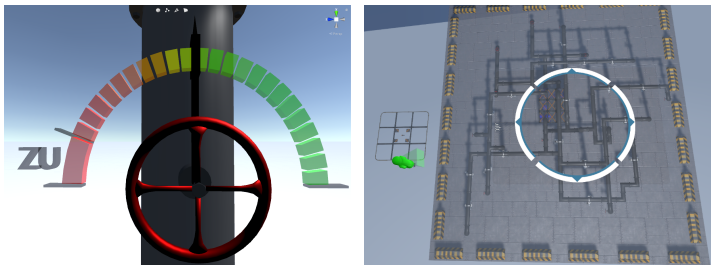
(b) Index finger touches the VirtualTablet as indicated by the green sphere.

Figure 5: Visualization of the VirtualTablet and the user's hand.

To visualize the VirtualTablet and the user's hands in VR, the engine Unity3D[5] is used. At the location of the real tablet, a virtual tablet with the same size and shape is rendered (see Figure 5). To give the user a visual feedback of the location of her or his hand, a 3D point cloud is rendered. White spheres at the fingertips highlight currently tracked fingers. The sphere turns green if a touch is detected.

The VirtualTablet can be used as normal tablet which is familiar to many users. However, VR allows to enhance and extend the interaction with the tablet. The following subsections describe different techniques that are enabled by VirtualTablet.

3.1 *Moving the VirtualTablet*



(a) A virtual valve wheel that can be turned by holding the VirtualTablet at the wheel's location and rotating it.

(b) The map of the VE can be moved by tilting the VirtualTablet. A green symbol represents the current location of the player. The user can teleport to a location by touching it.

Figure 6: Examples for a rotating tablet interaction.

Even without the touch recognition the VirtualTablet allows different forms of interaction. The position and rotation of the tablet can be used as an input to e.g. open a menu or change a value if held inside a trigger volume. If a user holds the tablet into a virtual object, an action can be triggered. By rotating the tablet at the location of a wheel, the wheel can be turned (see Figure 6a). We implemented a navigation mechanism that allows the user to move a map around by tilting the tablet (see Figure 6b). If the user touches a location she or he is teleported there.

Since the VirtualTablet has ArUco markers on both sides, it can recognize touch on both surfaces. This can be used to give the

[5]see https://unity3d.com/

user a more natural way to navigate through menus. If the user flips the tablet around, a different user interface is displayed. In our application it is used to switch between an information interface and a map display for navigation.

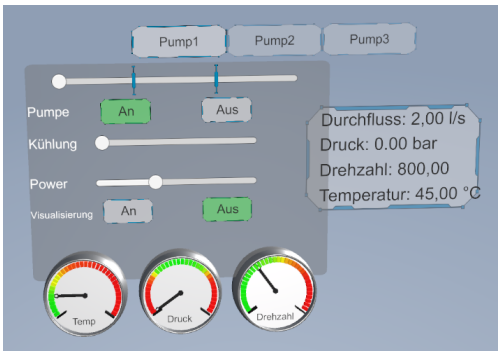
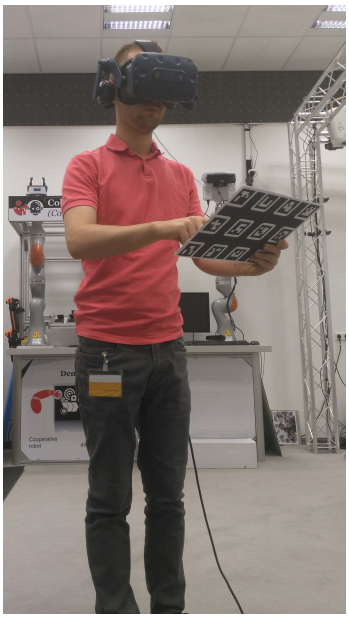


Figure 7: A menu that extends non-interactable information over the edges of the physical tablet.

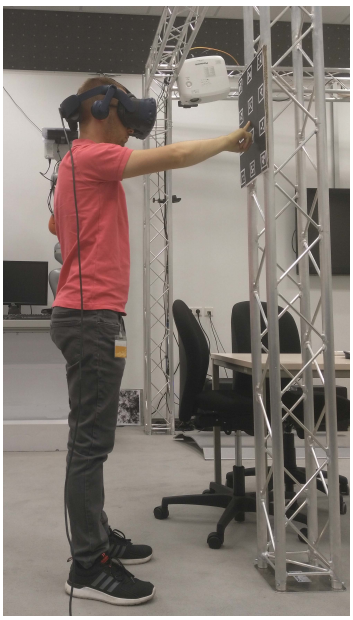
3.2 *Extending the VirtualTablet*

VR allows to display information at any given location. In reality, displays are much more constrained and information cannot (yet) be displayed in mid-air. The VirtualTablet allows to extend the displayed content over the edges of the physical tablet (see Figure 7). All interactable elements, like buttons and sliders, are displayed on the touchable surface. Other informations are arranged around the tablet.

3.3 *Duplicating the VirtualTablet*



(a) Mobile acrylic glass.



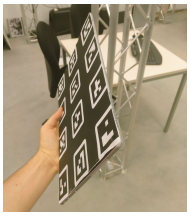
(b) Mounted cardboard.

Figure 8: Different interaction surfaces.

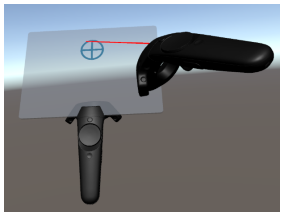
Physical tablets are expensive, rely on a power supply and only come in distinct shapes and sizes. With our approach multiple interactable surfaces can be created fast and cheap. As seen in Figure 8 we created a small mobile tablet and a larger mounted interaction surface. The different surfaces can be used in parallel.



(a) Attaching the ArUco board to the capacitive touchscreen tablet.



(b) Holding the capacitive touchscreen tablet.



(c) Laserpointer and interaction surface attached to the HTC Vive controllers.

Figure 9: Setup of the input techniques for comparison with the VirtualTablet.

4 Evaluation

To evaluate the proposed system and interaction methods, a user study was performed. We designed an application that allows a user to control different machines and valves in a factory using the VirtualTablet. The task of the user was to react to a breach in the pipe system. While solving the issues, the user interacted repeatedly with the extended interaction methods described above.

4.1 *Independent Variables*

To compare the detection performance of the VirtualTablet, other input methods were implemented. A capacitive touchscreen tablet was equipped with a ArUco board (see Figure 9a and 9b). The capacitive tablet uses the same hand visualization as the VirtualTablet, but the touch detection from the display as a ground truth. The size of the tablet is about 1 cm smaller in width and 0.5 cm smaller in height than the acrylic glass VirtualTablet. The default interaction tools for the HTC Vive are the provided controllers. The pose of the controllers is tracked with sub-millimeter accuracy [15]. Applications often use a laser pointer to interact with a handheld menu (see Figure 9c). This virtual pointer technique is effective and efficient [16]. Furthermore, a normal capacitive tablet in a non-VR scenario was used to evaluate the ground truth precision of touch interaction for the given tasks.

The user study was performed in a within-subjects design. The independent variables are the the four described techniques VirtualTablet, CapacitiveTablet and Controller (all in VR) as well as the NonVRTablet. To compensate for the effects of learning and fatigue the conditions were counterbalanced.

4.2 *Procedure*

First, the users tested the extended interaction methods in the factory application. Second, the users were asked to perform several click and draw interactions with the tablet as in [12]. Users were presented with 6 targets which appeared four times each per technique. Each target had a size of 5×5 cm, independent of the condition. As a result each user performed $6 \times 4 \times 4 = 96$ clicks in total. The order of the click targets was random, but pre-calculated for each

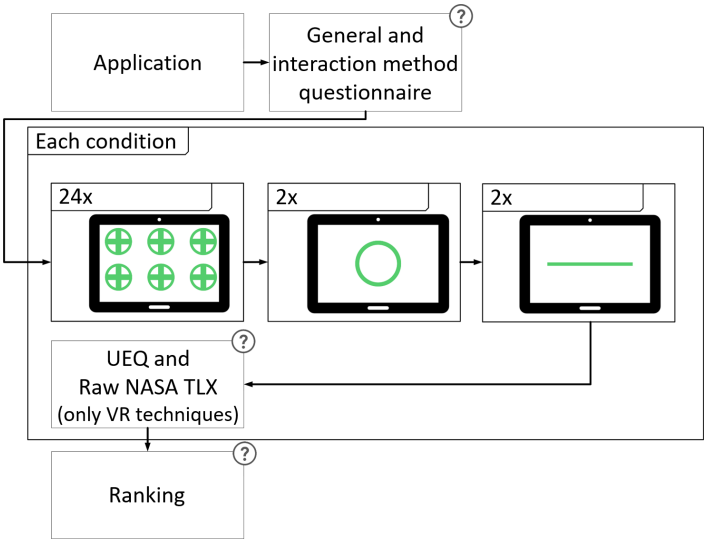


Figure 10: Procedure of the user study. Question mark indicates questionnaire.

technique so each user had the same sequence. For the shape tracing task, users were asked to trace a given figure (circle and line) as closely as possible and finishing their drawing by clicking a button. Each shape was drawn twice. The full procedure is listed in Figure 10.

4.3 Participants

28 people (4 female) participated in the user study with an age distribution of 2 under 20 years, 13 from 21-30 years, 8 from 31-40 years and 5 larger than 40 years. All of which had prior experience with touch input. On a scale of 1 (none) to 7 (very much), users had Ø 3.2 ± 1.9 experience with VR, 15 users had used a tracked controller before.

5 Analysis

5.1 Selection

The precision of the target selection was measured by recording the position of the touch and calculating the distance to the center of the target. We removed 49 outlier points (1.8%) with a distance of more than three standard deviations from the target point (as in [12], [17]). Also, the time taken was measured.

5.2 Tracing

For the shape tracing tasks, we measured how much of the target area was filled. Furthermore, the percentage of pixels painted inside the image was calculated. This value is an indication how accurately a user could draw inside the target area. Moreover, we calculated the average distance of the touch point to the target shape with the width of the brush and the shape removed. Also the duration of drawing was measured.

5.3 Performance and User Experience

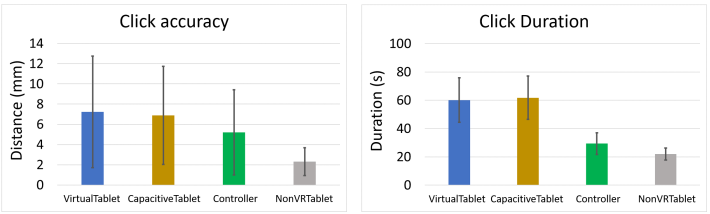
To assess the qualitative performance and experience of the different hardware setups, we used Raw NASA TLX [18] and the shortened version of UEQ, the UEQ-S [19] as a usability measurement. These questionnaires were only collected for the VR techniques since the NonVRTablet is only used as a ground truth for precision. At last, users were asked to rank the three VR techniques on a 7-point Likert scale regarding wearing comfort, quality of input and an overall ranking.

5.4 Statistical analysis

We present the results as average value (Ø) with standard deviation (±). We used the *rTOST* [20] as a equivalence test and the *Mann-Whitney-Test* [21] as a significance test. Both tests do not assume a normal distribution of the sample set. To calculate the power of the effects, *Cohen's d* [22] is used.

6 Results

6.1 Selection accuracy



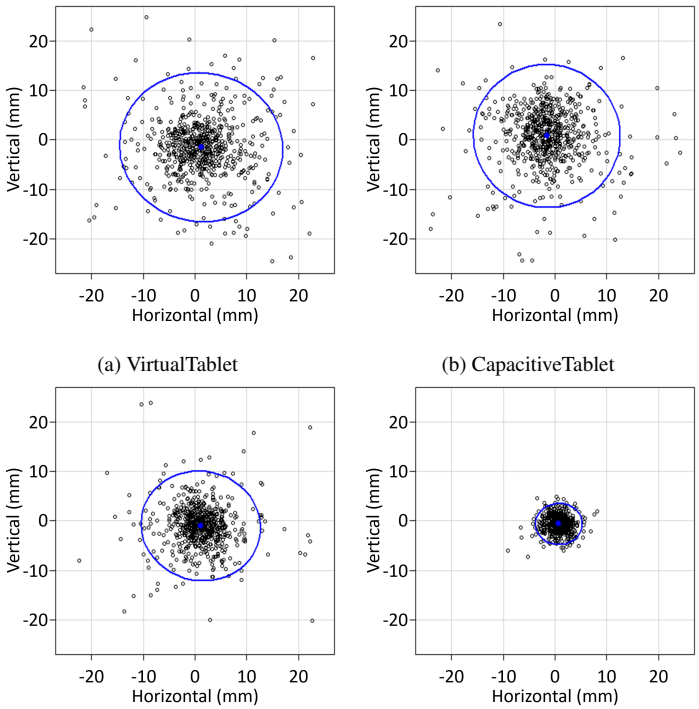
(a) Average selection distance in mm. (b) Average time for all 24 clicks.

Figure 11: Accuracy and duration of selections.

Table 1: Average selection distance in mm. The 6 targets are coded as T/B = top or bottom row and L/C/R = left, center or right column.

	Total	TL	TC	TR	BL	BC	BR
VirtualTablet							
Ø	6,877	6,528	5,401	7,155	8,005	7,272	6,910
±	4,869	4,701	3,603	4,827	5,287	5,287	4,985
CapacitiveTablet							
Ø	6,488	6,553	6,491	7,218	6,752	5,877	6,023
±	4,077	3,765	4,198	3,853	4,790	4,093	3,581
Controller							
Ø	4,827	5,028	5,039	5,221	4,905	4,162	4,614
±	3,291	3,445	2,929	3,619	3,345	2,783	3,501
NonVRTablet							
Ø	2,246	2,527	2,118	2,301	2,562	2,159	1,820
±	1,231	1,300	1,160	1,287	1,308	1,033	1,148

The results of the selection tasks are listed in Figure 11a and Table 1. The three VR input methods VirtualTablet, CapacitiveTablet



(a) VirtualTablet (b) CapacitiveTablet

(c) Controller (d) NonVRTablet

Figure 12: 95% confidence ellipses.

Table 2: rTOST equivalence test between pairs of VirtualTablet (VT), CapacitiveTablet (CT) and Controller (C)

	VT - C	VT - CT	CT - C
Epsilon	2,315	2,315	2,315
Avg. diff.	1,652	-0,001	-1,658
df	687,239	759,937	778,250
Cohen's d	0,412	0,065	0,370
p value	0,003	0,000	0,001

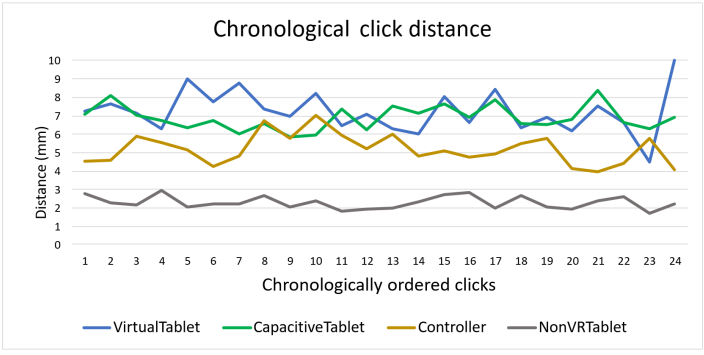


Figure 13: Average distance of clicks chronologically ordered.

and Controller have a selection accuracy of about 5 to 7 mm. The 95% confidence ellipses in Figure 12 show visually, what the rTOST test (see Table 2) confirms. All three VR input methods are equivalent. The epsilon of 2.315 is the average distance of the ground truth tablet interaction, meaning this value represents the accuracy of the users in a best case scenario. Analysis of the accuracy of the

Table 3: Average duration for 24 selections.

	VirtualTablet	CapacitiveTablet	Controller	NonVRTablet
Ø	60,120	61,770	29,350	22,060
±	15,690	15,270	7,695	4,243

6 targets as seen in Figure 1 reveals, that the VirtualTablet detection has a higher accuracy of about 1 mm at the top row. The other techniques do not show such a large difference. Also, no learning effect can be detected in the collected data. Figure 13 shows the average distances to the center in a temporal order. The difference from the center stays approximately the same over the period of the selection task. The total amount of time needed to perform the selection task with each input technique is shown in Figure 11b and Table 3. Users took about twice as much time with the VirtualTablet and CapacitiveTablet in regards to the Controller interaction and almost three time as much in comparison to the NonVRTablet.

6.2 *Tracing accuracy*

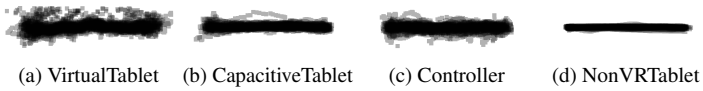


(a) VirtualTablet (b) CapacitiveTablet (c) Controller (d) NonVRTablet

Figure 14: Stacked drawings of all circle traces.

Table 4: Shape tracing accuracy for the circle.

	VirtualTablet	CapacitiveTablet	Controller	NonVRTablet
	% of drawing inside target area			
Ø	35,131	47,709	45,059	63,884
±	8,884	11,847	11,239	8,949
	% of target area filled			
Ø	56,840	66,057	63,671	77,615
±	14,452	16,490	17,694	13,091
	Average distance to target shape (mm)			
Ø	4,274	2,960	2,892	1,873
±	2,106	1,851	1,003	0,590
	Drawing duration (s)			
Ø	15,688	8,128	6,455	5,277
±	8,540	2,877	3,542	2,378



(a) VirtualTablet (b) CapacitiveTablet (c) Controller (d) NonVRTablet

Figure 15: Stacked drawings of all line traces.

Table 5: Shape tracing accuracy for the line segment.

	VirtualTablet	CapacitiveTablet	Controller	NonVRTablet
	% of drawing inside target area			
Ø	39,058	56,813	49,715	82,004
±	9,464	14,805	18,014	11,840
	% of target area filled			
Ø	61,949	67,469	64,202	78,607
±	13,468	17,716	23,724	11,706
	Average distance to target shape (mm)			
Ø	4,340	2,757	3,430	2,400
±	1,383	1,387	1,640	0,848
	Drawing duration (s)			
Ø	6,885	3,909	3,148	2,521
±	4,075	1,424	2,179	1,172

The results for the shape tracing task are shown in Table 4 and 5 for the circle target and line target respectively. Drawing with the NonVRTablet achieves the highest tracing accuracy regarding percentage of drawing inside the target area and target area filled. The CapacitiveTablet is on average slightly better than the Controller and VirtualTablet condition. The standard deviation is smallest for the VirtualTablet and the NonVRTablet. The average distance towards the circle target area is very similar to this, with the NonVRTablet performing best with 1.9 mm accuracy and VirtualTablet performing worst with a distance of 4.3 mm. For the line segment, the average distance of the CapacitiveTablet is very close to the NonVRTablet with 2.8 mm to 2.4 mm. The detected touches of the VirtualTablet have the highest average distance (4.3 mm), but are not so far away from the Controller tracing (3.4 mm). Figures 14 and 15 show the shape traces of all users on top of each other. The touch detection of the VirtualTablet leads to a lot of smaller errors outside of the target area. The ground truth touch of the CapacitiveTablet is a lot more stable. The Controller drawings appear shaky. The NonVRTablet matches the target area the most.

Regarding shape tracing duration, the Controller is on average faster, but more inaccurate than the CapacitiveTablet. The NonVRTablet is the quickest and the VirtualTablet takes the most time.

6.3 Performance and User Experience

The results of the evaluation show that the arrangement of the extended menus with information displayed outside the interactable area was clear (see Table 6). It was also very useful to have more than one input surface. Teleportation with the tilting map has a medium difficulty. During the user study users often needed help to initially understand what they needed to do. Yet, once learned, participants quickly got better. Turning the valves with the orientation of the tablet and using the displayed buttons and sliders also has a medium difficulty. The pose of the tablet and touch input is recognized medium-well. The participants are happy with the interaction distance, which is due to the camera sensor at least 30 cm in front of the HMD. Input delay was quite high (mostly due to the cameras), which is a little bit annoying for the users. However, the haptic surface for hand gesture input is clearly rated as helpful.

Table 6: 7 point Likert scale questionnaire for different aspects of the extended interaction techniques of the VirtualTablet.

Questions	Value 1	Value 7	Ø	±
Information outside the input area was...	unclear	clear	5,143	1,597
Using more than one input surface are was...	practical	impractical	2,429	1,116
Teleportation with the tilting map was...	easy	difficult	4,107	1,718
Using the tablet orientation as input was...	easy	difficult	3,607	1,800
Using the buttons and sliders was...	easy	difficult	3,643	1,315
Reaction of touch input was was...	expected	unexpected	3,786	1,319
Recognition of tablet pose was...	expected	unexpected	2,571	1,400
Minimum distance for input was...	too far away	just right	5,071	1,534
Input delay was...	annoying	undisturbing	3,786	1,839
Haptic surface for input was...	helpful	unnecessary	1,500	1,086

Figure 16 shows the results of the UEQ-S ratings. The pragmatic quality is very good for the Controller, positive for the CapacitiveTablet and neutral for the VirtualTablet (positive evaluation from a value of 0.8 and more). the hedonic quality is rated positive for all techniques, with the VirtualTablet and Controller rated best. The overall result shows that the Controller technique works very well. However, both touch surface interactions receive a positive rating. Compared to the supplied benchmark ratings from the UEQ Data Analysis Tool [23] the VirtualTablet is ranked as bad for the pragmatic quality, above average for the hedonic quality and below average overall.

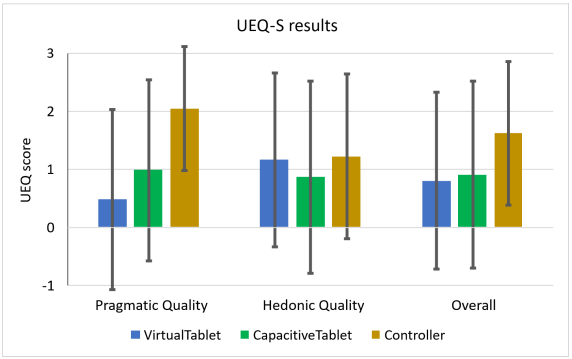


Figure 16: Results of UEQ-S.

The CapacitiveTablet is below average for all three categories. The Controllers are ranked as excellent for the pragmatic quality and overall. Their hedonic quality is above average. The difference between the Controller technique and the two touch techniques is significant for the pragmatic quality ($p \leq 0.001$) and overall

($p \leq 0.002$ and $p \leq 0.012$ for the difference towards the VirtualTablet and CapacitiveTablet respectively).

The results of the Raw NASA TLX as seen in Figure 17 show that the Controller ranks best in all categories. The VirtualTablet ranks worst in almost all categories, except for the physical demand.

Users rank the wearing comfort of the VirtualTablet and the Controllers as good, the CapacitiveTablet as below medium (see Figure 18). The quality of input is ranked best for the Controllers, good for the CapacitiveTablet and below medium for the VirtualTablet. Overall, the participants rank the Controller as good and the tablet techniques as medium.

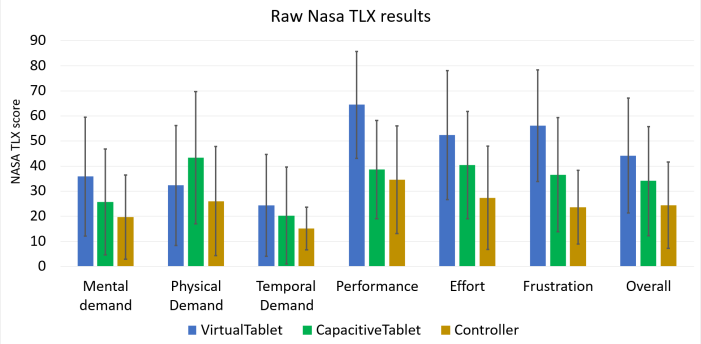


Figure 17: Results of Raw NASA TLX.

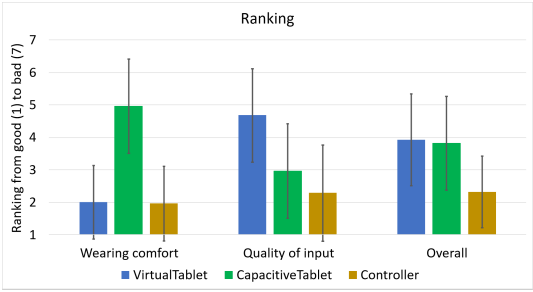


Figure 18: Rankings of the techniques.

7 Discussion

The rTOST test shows that both the VirtualTablet and the Controller interaction perform equally in a selection task. The precision of the VirtualTablet is about 6.9 mm, which is over half as small as the average size of a fingertip with 1620 mm in diameter [24]. The precision is good enough to interact with objects of suggested minimum target sizes, e.g. 9.6 mm [25] or 7-10 mm[6]. Because of the larger temporal delay, the tablet interaction is slower. This yields to longer interaction times and also affects the precision of the task execution, since the visualized hand and tablet locations did not match current location of their real world counterpart. The interaction with a controller is different compared to touch interaction. A user often only twists the wrist to point the laser towards another target. This is quicker than moving the whole hand, but it also induces more jitter as the larger standard deviations show. The shape tracing task shows that the touch detection of VirtualTablet is not robust enough. Due to the large amount of noise and invalid data in the calculated depth information, the touch detection is not able to always detect a continuous touch and sometimes detects false positives. The touch detection also fails, if a user obscures the fingertip with her or his own hand. This explains the larger time difference and usability issues of the VirtualTablet compared to the CapacitiveTablet and Controller techniques, which recognize the touch through sensors.

VR does not allow the user to see her or his own hand. A good visualization of the user's hand is therefore necessary to achieve good results. The used point cloud and fingertip spheres depend on the segmentation of the hand. The blur during head movements and the automatic brightness adjustments dampen the quality of the image. However, the point cloud representation by itself seems to be not so easy to understand spatially. The impact of the hand visualization in the shape tracing task is lower than in the selection task for the CapacitiveTablet, because the user receives direct feedback from her or his touch point on the capacitive display and can compensate for any hand visualization delay or errors. Thus, the CapacitiveTablet performs better than the VirtualTablet.

The haptic surface for the interaction was rated as very helpful during the application task. This was also indicated through comments by the participants. The acrylic glass material is very lightweight (120 g) compared to the touchscreen tablet with battery (754 g) which leads to a lower physical demand and a higher wearing comfort (Vive controllers 2×203 g). The impact of the weight will be even larger at longer sessions of usage.

The comparable MRTouch system [12] shows slightly better selection precision (Ø 5.4 ± 3.2 mm) when compared to the VirtualTablet. However, the VirtualTablet interaction uses movable touch surfaces, which induce additional precision erros. Furthermore, because MRTouch is an AR augmented reality (system, there is no need for a hand or surface visualization. But, the visualization in VR is crucial and has a larger impact on the precision and usability of the input technique. Tracing accuracy of VirtualTablet and MRTouch are comparable with an average distance of Ø 4.0 ± 3.4 mm for MRTouch. Although the AR HMD of MRTouch uses a time of flight depth sensor and infrared cameras, the touch detection works with a threshold distance of 10 mm (ours 5 mm) which could lead to a touch recognition before the user reached the surface. The very accurately tracked Controllers show a similar target selection result as VirtualTablet and MRTouch. This shows that touch input in VR and AR is very accurate, even with the present issues regarding inaccurate tracking.

Our ground truth baseline (NonVRTablet with Ø 2.3 ± 1.4 mm avg. distance to target) shows that there is room for improving the accuracy of the presented technique. Yet, the extended functionality, e.g. extending the user interface over the edges of the physical surface, turning the tablet around for a menu switch or changing the pose of the tablet to control a device, has a benefit that a real-world touch tablet cannot offer.

8 Conclusion and Future Work

We presented VirtualTablet, a movable surface that allows familiar touch interaction to control the VE. VirtualTablet is intuitive to

[6]see https://support.google.com/accessibility/android/answer/7101858

control, cheap, lightweight and can be shaped into any form and size. It can be used as an extension to currently supported input hardware, like controllers, since it does not require any additional hardware setup (except the board itself) and does not alter the VR HMD in any way. The haptic surface is helpful and the implementation allows an accurate selection of targets which is comparable to a laserpointer-based interaction using controllers.

The hardware used in this study has high latency and large noise levels. This yields to issues with continuous touch detection and dampens the overall accuracy of the system. For future work we would like to use a different camera and/or depth sensor to resolve these challenges. Also, other sensors like a capacitive touch foil on the tablet could be used to eliminate the problems of occlusion. We proposed several new interaction techniques that extend the usefulness of touch interaction in VR environments. With better tracking, further exploration of the design space of touch in VR will be possible. Another aspect that we need to improve is the hand visualization so that users get a better understanding of their finger positions. Mesh- or skeleton-like representations could be used for this purpose.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment We want to thank all of the participants of the user study for taking the time to provide us with very helpful feedback and insights.

References

[1] A. H. Hoppe, F. Marek, F. van de Camp, and R. Stiefelhagen, "Virtualtablet: Extending movable surfaces with touch interaction", in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Mar. 2019, pp. 980–981. DOI: 10.1109/VR.2019.8797993.

[2] D. Markov-Vetter, E. Moll, and O. Staadt, "Evaluation of 3d selection tasks in parabolic flight conditions: Pointing task in augmented reality user interfaces", in *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, ser. VRCAI 12, Singapore, Singapore: Association for Computing Machinery, 2012, 287294, ISBN: 9781450318259. DOI: 10.1145/2407516.2407583.

[3] E. Koskinen, T. Kaaresoja, and P. Laitinen, "Feel-good touch: Finding the most pleasant tactile feedback for a mobile touch screen button", in *Proceedings of the 10th International Conference on Multimodal Interfaces*, ser. ICMI 08, Chania, Crete, Greece: Association for Computing Machinery, 2008, 297304, ISBN: 9781605581989. DOI: 10.1145/1452392.1452453.

[4] C. S. Tzafestas, K. Birbas, Y. Koumpouros, and D. Christopoulos, "Pilot evaluation study of a virtual paracentesis simulator for skill training and assessment: The beneficial effect of haptic display", *Presence*, vol. 17, no. 2, pp. 212–229, Apr. 2008, ISSN: 1054-7460. DOI: 10.1162/pres.17.2.212.

[5] S. Scheggi, L. Meli, C. Pacchierotti, and D. Prattichizzo, "Touch the virtual reality: Using the leap motion controller for hand tracking and wearable tactile devices for immersive haptic rendering", in *ACM SIGGRAPH 2015 Posters*, ser. SIGGRAPH 15, Los Angeles, California: Association for Computing Machinery, 2015, ISBN: 9781450336321. DOI: 10.1145/2787626.2792651.

[6] A. L. Simeone, E. Velloso, and H. Gellersen, "Substitutional reality: Using the physical environment to design virtual reality experiences", in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI 15, Seoul, Republic of Korea: Association for Computing Machinery, 2015, 33073316, ISBN: 9781450331456. DOI: 10.1145/2702123.2702389.

[7] B. Araujo, R. Jota, V. Perumal, J. X. Yao, K. Singh, and D. Wigdor, "Snake charmer: Physically enabling virtual objects", in *Proceedings of the TEI 16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, ser. TEI 16, Eindhoven, Netherlands: Association for Computing Machinery, 2016, 218226, ISBN: 9781450335829. DOI: 10.1145/2839462.2839484.

[8] L. Kohli, "Redirected touching: Warping space to remap passive haptics", in *2010 IEEE Symposium on 3D User Interfaces (3DUI)*, Mar. 2010, pp. 129–130. DOI: 10.1109/3DUI.2010.5444703.

[9] M. Azmandian, M. Hancock, H. Benko, E. Ofek, and A. D. Wilson, "Haptic retargeting: Dynamic repurposing of passive haptics for enhanced virtual reality experiences", in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI 16, San Jose, California, USA: Association for Computing Machinery, 2016, 19681979, ISBN: 9781450333627. DOI: 10.1145/2858036.2858226.

[10] L.-P. Cheng, E. Ofek, C. Holz, H. Benko, and A. D. Wilson, "Sparse haptic proxy: Touch feedback in virtual environments using a general passive prop", in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI 17, Denver, Colorado, USA: Association for Computing Machinery, 2017, 37183728, ISBN: 9781450346559. DOI: 10.1145/3025453.3025753.

[11] D. Medeiros, L. Teixeira, F. Carvalho, I. Santos, and A. Raposo, "A tablet-based 3d interaction tool for virtual engineering environments", in *Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, ser. VRCAI 13, Hong Kong, Hong Kong: Association for Computing Machinery, 2013, 211218, ISBN: 9781450325905. DOI: 10.1145/2534329.2534349.

[12] R. Xiao, J. Schwarz, N. Throm, A. D. Wilson, and H. Benko, "Mrtouch: Adding touch input to head-mounted mixed reality", *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 4, pp. 1653–1660, Apr. 2018, ISSN: 2160-9306. DOI: 10.1109/TVCG.2018.2794222.

[13] S. Garrido-Jurado, R. Muoz-Salinas, F. Madrid-Cuevas, and M. Marn-Jimnez, "Automatic generation and detection of highly reliable fiducial markers under occlusion", *Pattern Recognition*, vol. 47, no. 6, pp. 2280 –2292, 2014, ISSN: 0031-3203. DOI: 10.1016/j.patcog.2014.01.005.

[14] N. S. Chethana, Divyaprabha, and M. Z. Kurian, "Design and implementation of static hand gesture recognition system for device control", in *Emerging Research in Computing, Information, Communication and Applications*, N. R. Shetty, N. H. Prasad, and N. Nalini, Eds., Singapore: Springer Singapore, 2016, pp. 589–596, ISBN: 978-981-10-0287-8.

[15] D. C. Niehorster, L. Li, and M. Lappe, "The accuracy and precision of position and orientation tracking in the htc vive virtual reality system for scientific research", *i-Perception*, vol. 8, no. 3, 2017. DOI: 10.1177/2041669517708205.

[16] I. Poupyrev, T. Ichikawa, S. Weghorst, and M. Billinghurst, "Egocentric object manipulation in virtual environments: Empirical evaluation of interaction techniques", *Computer Graphics Forum*, vol. 17, no. 3, pp. 41–52, 1998. DOI: 10.1111/1467-8659.00252.

[17] C. Harrison, H. Benko, and A. D. Wilson, "Omnitouch: Wearable multitouch interaction everywhere", in *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, ser. UIST 11, Santa Barbara, California, USA: Association for Computing Machinery, 2011, 441450, ISBN: 9781450307161. DOI: 10.1145/2047196.2047255.

[18] S. G. Hart, "Nasa-task load index (nasa-tlx); 20 years later", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 9, pp. 904–908, 2006. DOI: 10.1177/154193120605000909.

[19] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Design and evaluation of a short version of the user experience questionnaire (ueq-s)", *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, p. 103, Jan. 2017. DOI: 10.9781/ijimai.2017.09.001.

[20] K. K. YUEN, "The two-sample trimmed t for unequal population variances", *Biometrika*, vol. 61, no. 1, pp. 165–170, Apr. 1974, ISSN: 0006-3444. DOI: 10.1093/biomet/61.1.165.

[21] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other", *The Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 50–60, 1947, ISSN: 00034851. DOI: 10.1214/aoms/1177730491.

[22] J. Cohen, *Statistical power analysis for the behavioral sciences. (2nd ed.)* 1988.

[23] A. Hinderks, M. Schrepp, and J. Thomaschewski, "A benchmark for the short version of the user experience questionnaire", Sep. 2018. DOI: 10.5220/0007188303730377.

[24] K. Dandekar, B. I. Raju, and M. A. Srinivasan, "3-D Finite-Element Models of Human and Monkey Fingertips to Investigate the Mechanics of Tactile Sense ", *Journal of Biomechanical Engineering*, vol. 125, no. 5, pp. 682–691, Oct. 2003, ISSN: 0148-0731. DOI: 10.1115/1.1613673.

[25] P. Parhi, A. K. Karlson, and B. B. Bederson, "Target size study for one-handed thumb use on small touchscreen devices", in *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI 06, Helsinki, Finland: Association for Computing Machinery, 2006, 203210, ISBN: 1595933905. DOI: 10.1145/1152215.1152260.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 334-339 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Skin Tissue Oxygen Saturation Prediction: A Comparison Study of Artificial Intelligence Techniques

Audrey Huong[1,*], Xavier Ngu[2]

[1]*Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Johor, Malaysia*

[2]*Centre for Applied Electromagnetics, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Johor, Malaysia*

ARTICLE INFO

Article history:
Received: 11 November, 2019
Accepted: 27 December, 2019
Online: 07 February, 2020

Keywords:
Skin tissue oxygen
Partial Least Square
K-Nearest Neighbor
Neural Network

ABSTRACT

Noninvasive measurement of skin tissue oxygen saturation, S_tO_2, is of interest especially in the studies of wound healing and detection of vascular diseases. This work aims to compare Partial Least Square (PLS) regression, K-Nearest Neighbor (KNN) and Artificial Neural Network (ANN) technique in the prediction of S_tO_2 using spectral data obtained from Monte Carlo simulations. We found a good performance of PLS technique with mean and standard deviation (SD) of the absolute prediction errors given by 2(1.37) % for noisy data with signal-to-noise ratio (SNR) of 30 dB. The errors increased to 4.67(3.87) % under SNR of 20 dB. This is followed by KNN technique, which calculated errors varied from 4.28(3.58) % to 7.61(6.08) % with a decrease in SNR. Meanwhile ANN produced large errors in its prediction ranging from 4 to 9 % for the considered SNR values. This work concluded PLS is suitable for use in selecting significant spectral information for the development of prediction models using artificial intelligent (AI) system or a hybrid combination of AIs to achieve a higher accuracy.

1. Introduction

Tissue oxygen saturation is a clinical indicator for assessment and diagnosis of many illnesses such as peripheral vascular diseases [1], organ viability [2], hypoxic-ischemia conditions [3], hemorrhagic shock [4], in studies related to cancers [5] and muscle oxidative metabolism [6], and for management of wounds (burns) and ulcers [5, 7]. In this context, sufficient tissue oxygen pressure combined with adequate supply of oxygen and nutrients through functioning microcirculatory system are more likely to guarantee cell survival.

The current state of the arts of skin tissue oxygen consumption measurement includes transcutaneous oxygen tension ($TcPO_2$). Despite its noninvasive and high accuracy attributes, this equipment required mild heating at the local site in its operation to promote blood perfusion prior to the measurement. Some of the other drawbacks include long measurement time, and sensitivity of its performance to factors such as the person's stress level, body temperature, medication and local inflammation [8]. Systems that take advantage of physical and optical properties of blood for noninvasive and noncontact measurement of blood oxygen saturation are of increasing demand; the said systems included spectroscopy techniques (optical, time and frequency domains) and clinical imaging methods such as Magnetic Resonance Imaging [9,10], and Photoacoustic techniques [11]. Among the spectroscopy approaches, hyperspectral imaging technique has emerged as a popular choice owing to its inherent ability to produce high resolution data in both spectral and spatial dimensions. Even though this system collects images at a shallow depth of field, it involves less complex technologies. With regard to its application in tissue oxygen measurement using light of visible range, it is able to spectrally resolve information carrying light backscattered from cutaneous tissue. The latter can be used in prediction of haemoglobin concentration and tissue oxygen saturation using either an inverse model, or computer aided, or computational intelligence system [12, 13]. Since the system interrogates blood in dermal capillaries, arterioles and venules [14], the predicted value is often taken as the average of oxygen saturation within the cutaneous microenvironment, S_tO_2. Previous works [15, 16] from this laboratory found that the range of predictions is dependent on the system configuration and architecture, and the wavelengths used. Meanwhile the use of Artificial Intelligence (AI) techniques has received a great deal of attention in this arena of study over the past years due to its ability to learn by itself through a large set of data, events and statistics before making decision. This capacity makes it a highly sought-

*Audrey Huong, Email: audrey@uthm.edu.my

after means for classification, prediction and forecasting tasks in many fields (for a review, see [17, 18]). In addition, much less efforts and time is needed to spend on improving the formulations of inverse models and look-up table.

The use of AI in medicine is well recognized to provide better quality medical health care services by reinforcing diagnosis with prediction and identifying definitive intervention, thus reducing morbidity and mortality case, and healthcare cost. Among the applications of AI include diagnosis and prognosis of cancers, illness and disease, to aid in decision makings and for physiotherapy practice [19, 20]. The examples of AI used for clinical screening and imaging tasks are such as optimization algorithms (see ref. [21] for a comprehensive review), neural network [22], and computer aided diagnosis [23].

To the best of the authors' knowledge the use of AI, and the study of the performance of different AI techniques, in the predictions of S_tO_2 using an optical system is few and far between. The further advancement in the integration of these technologies is impeded by insufficiency of data available for the training of the AI system, prompting the work by [24] that used Monte Carlo (MC) platform for training of artificial neural networks for future optical diagnostics and sensing. This work is following in the footsteps of [24], and our aim is to compare the performance of different AI techniques for the prediction of S_tO_2 using MC simulation data.

2. Material and Methods

2.1. Monte Carlo (MC) simulations of light propagation

The two layered human skin model, the illumination and detection systems used for simulations of light propagated in skin shown in Figure 1 were previously presented in [22]. These simulation data are used here again owing to the known medium S_tO_2 for evaluations of performance of prediction techniques described in the following. The simulation data for medium percent S_tO_2 ranging from 0 % to 100 % at a step resolution of 5 % were added with Gaussian White Noise (AWGN function in MATLAB 2018a) to produce signals with signal-to-noise ratio (SNR) of 20 and 30 dB. For each SNR and S_tO_2 level, 10 sets of random noise were incorporated into the simulations to give a total of 210 data for each SNR level, among which 70 % (n = 147) were randomly chosen for the training of the model, while the remaining (n = 63) were used as the testing data.

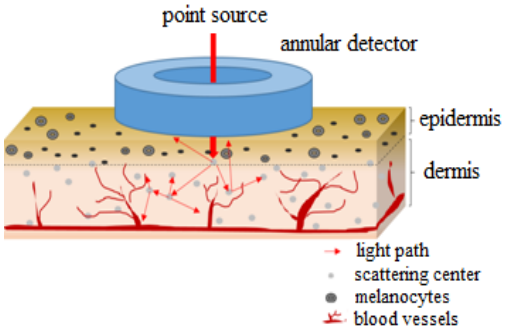


Figure 1: MC simulations of light propagation in a two-layered skin medium

2.2. Features selection

Since light absorption in skin varied with oxygen binding status of hemoglobins (i.e. oxyhemoglobin and deoxyhemoglobin concentration), by examining the absorption differences in the spectra of these hemoglobin components one can deduce the required S_tO_2 parameter. The simulated spectral data in section 2.1 were used as the input data to the artificial intelligence techniques. This work considered signals across the wavelengths ranged from 520 nm to 650 nm where changes in hemoglobin absorptivity signatures with S_tO_2 are significant (image can be found in [25]).

2.3. Tissue oxygen saturation prediction tools

In the present study, we investigated three different machine learning techniques: artificial neural network (ANN), partial least square (PLS) and K-nearest neighbor (KNN) in the estimation of the required percent S_tO_2 value. The error, ε, (in unit of %) in the predicted value, $S_tO_{2(p)}$, is defined as:

$$\varepsilon = S_tO_{2(p)} - S_tO_2' \quad (1)$$

where S_tO_2' is the ground truth value. The prediction tools presented in the following used MATLAB (version 2018b) software in their implementation due to its comprehensive multivariate analysis toolboxes.

2.3.1. Artificial Neural Network (ANN)

This work used *nnstart* toolbox to generate an ANN feedforward model from the simulation data. We have chosen 10 hidden neuron layers, and the fitting of the input-output relationship was by using Levenberg-Marquardt (LM) backpropagation model to update weights in minimizing sum of square error functions. This ANN model was modified from the prior work in [22], wherein 'while' loop structure was adopted in obtaining a neural network with acceptable performance during its training. This is with the implementation of arbitrarily chosen MSE threshold value of 24 and 15, respectively, for dataset of SNR = 20 and 30 dB.

2.3.2. Partial Least Square Regression (PLS)

PLS regression analysis was used here to develop a prediction model. This technique predicts the response, Y, for a set of data by extracting latent variables from sampled observed variables, X, as followed:

$$\begin{aligned} X &= T * P' + E_0 \\ Y &= U * Q' + F_0 \end{aligned} \quad (2)$$

where T and U are latent variables, P' and Q' are loading matrices and E_0 and F_0 are residuals. The latent variables are related via inner relation coefficient, B, and residual, F_1, as followed:

$$U = TB + F_1. \quad (3)$$

This study used *plsregress* function to project high dimensionalities feature vectors (X and Y) onto a subspace to give an overall partial regression model given by:

$$\tilde{Y} = X * (P * B * Q') + F. \quad (4)$$

The $P * B * Q'$ vector in (4) that simplified the regression problem was used to predict the most probable percent S_tO_2 value.

We have chosen latent factor of 3 for demonstration purpose. It must also be mentioned that the present work found no improvement in the prediction accuracy using other factor number.

2.3.3. K-Nearest Neighbor (KNN)

The KNN classification model for predictors, X, and response, Y, was generated using *fitcknn*. This study considered exhaustive search algorithm to find the nearest neighbors by calculating the distances from all points in X to each point in Y. The distance metric that was considered in the present study is Minkowski distance. The K value was chosen as 11 based on the size of the datasets following the discussion in [26]. The trained classifier in (5) was then used to predict the class of the testing data.

$$mdl = fitcknn\,(X, Y). \tag{5}$$

3. Results and analysis

The changes in light attenuation spectrum following variations in the reference S_tO_2 (i.e. ground truth values used in the simulations) are shown in Figure 2. The plots showed example changes in the noise corrupted signals of the considered SNR. These spectral data across wavelength range of 520 – 650 nm and a resolution of 1 nm were used for both training and evaluation of the considered models using a constant random seed for comparable results. The data were separated into training and testing sets with a percentage of 70:30. Figure 3 showed the three-dimensional plots of attenuations from training set for the reference of our readers. These plots also revealed changes in the value range and distribution with noise level. The mean (SD) of the absolute ε calculated from (1) for percent value predicted by ANN, PLS and KNN technique for spectrum of different reference S_tO_2 shown in Table 1 is calculated as 8.76 (6.82) %, 4.67 (3.88) % and 7.62 (6.08) %, and 4.77 (3.96) %, 2 (1.37) % and 4.28(3.57) % for dataset of SNR of 20 dB and 30 dB, respectively. The computation time (averaged from three consecutive runs) spent in the training of the machine learning techniques described in section 2.3.1-2.3.3 performed on an ASUS laptop with 64-bit window 7, Intel® Core™ Duo T6500 CPU @2.10 GHz are given by 1.38 and 1.8 seconds for PLS and KNN, respectively. A longer average time was needed for ANN with the time taken recorded as 81.5 and 39.1 seconds in the training of data of SNR = 20 and 30 dB, respectively.

4. Discussion

Figure 2 showed substantial decrease in the noise level on the spectra for signals of SNR 30 dB as compared to that of 20 dB. This decrease in noise level yields an overall improvement in the prediction accuracy in Table 1 with an average decrease in ε value calculated as 3.33(**0.66**) %. The lowest mean prediction error was observed for values given by PLS technique; this is followed by KNN and ANN. While both ANN and PLS produced regressed value in their predictions, categorical values according to the range of target values used during the training process were produced by the KNN. The main reason PLS technique outperformed its counterparts in the validation phase is likely due to the accentuation of the variations in the absorptivity through the projection of data to a lower dimensional space. An investigation into the correlation coefficients obtained from the regression of centre variables X and Y in (2) using three latent factors revealed a strong positive correlation at wavelength 580 nm, followed by that of 538 nm under SNR = 30 dB. These peaks increasingly diminished under SNR 20 dB.

Since similar distribution of attenuation values with changes in S_tO_2 was observed in Figure 2, and the features used in the training of the considered AI models discussed in section 2.2 are of full spectral information (attenuations across wavelength range 520 –

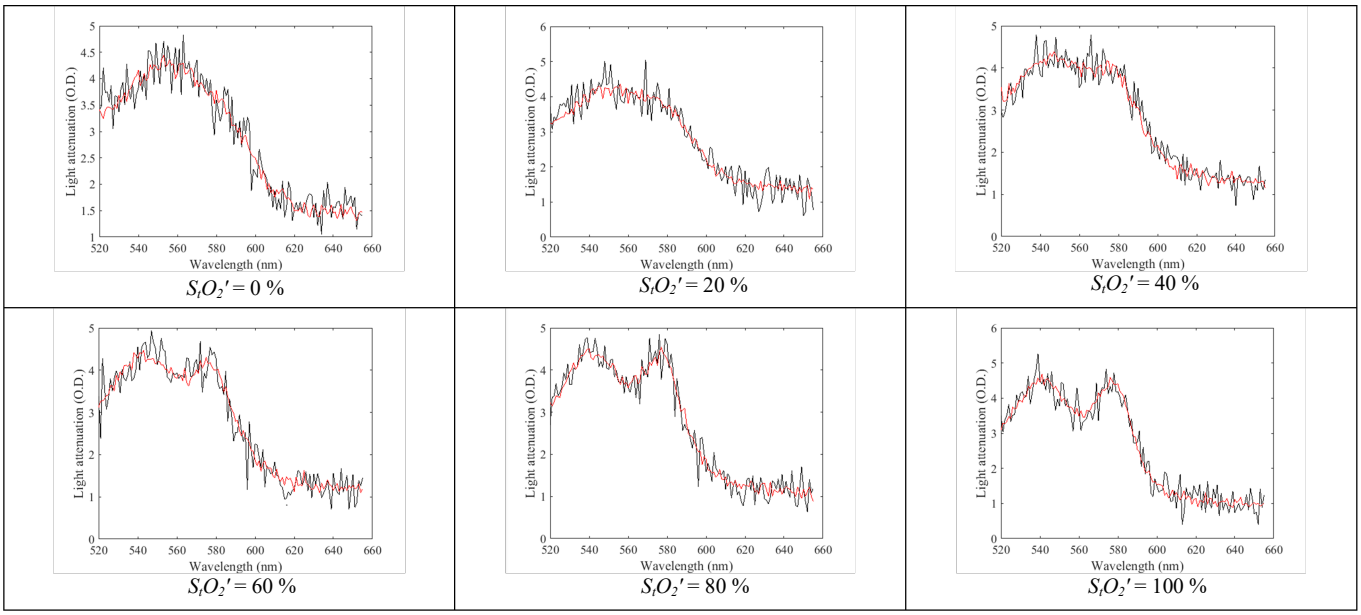
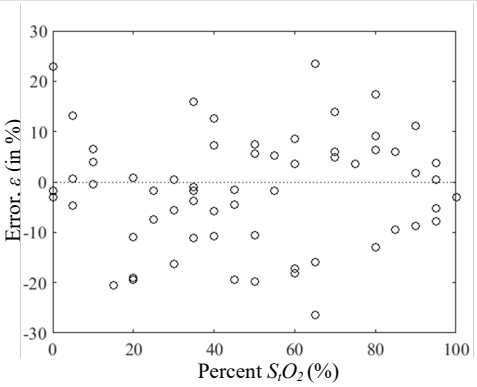
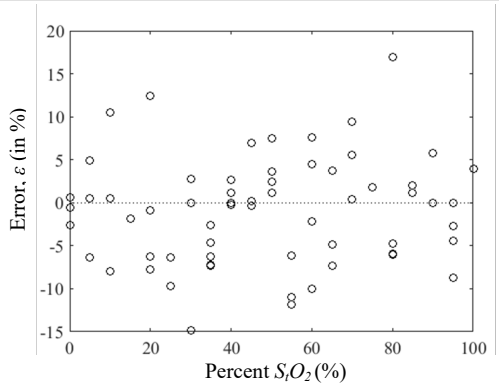
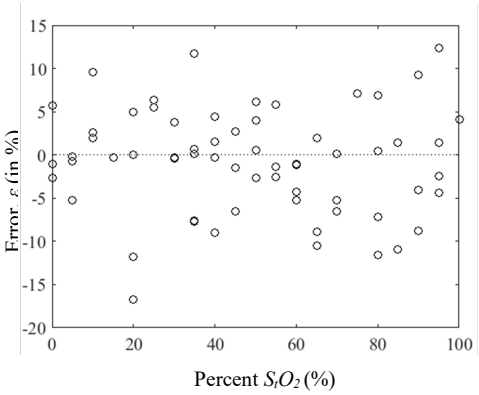
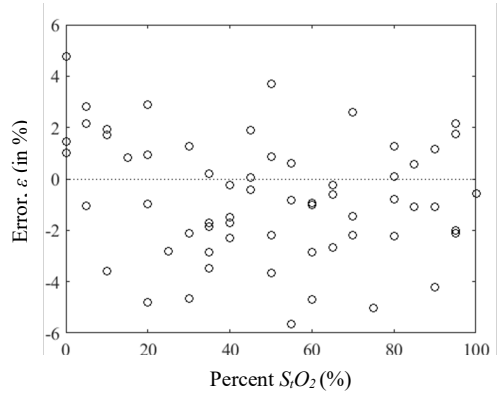
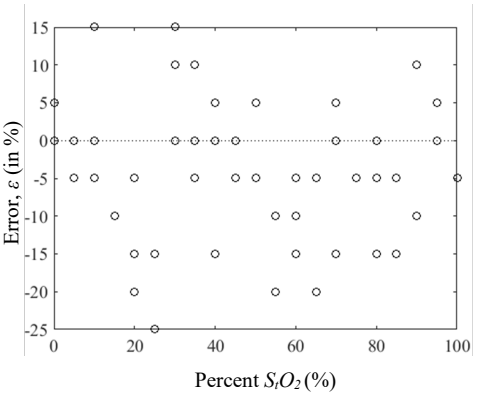
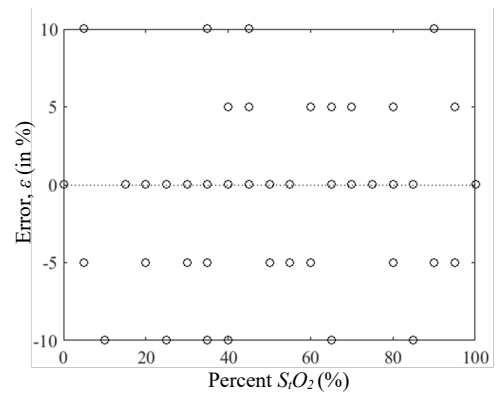


Figure 2: Noise corrupted wavelength dependent light attenuation with Signal to Noise (SNR) of 20 dB (dark lines) and 30 dB (red lines) under different S_tO_2 condition.

Table 1: Error, ε, in the value predicted for testing dataset of SNR 20 dB (Left) and 30 dB (Right) using (from top to bottom) Artificial neural network (ANN), Partial Least Square (PLS) and K-Nearest Neighbor (KNN).

ANN		
PLS		
KNN		

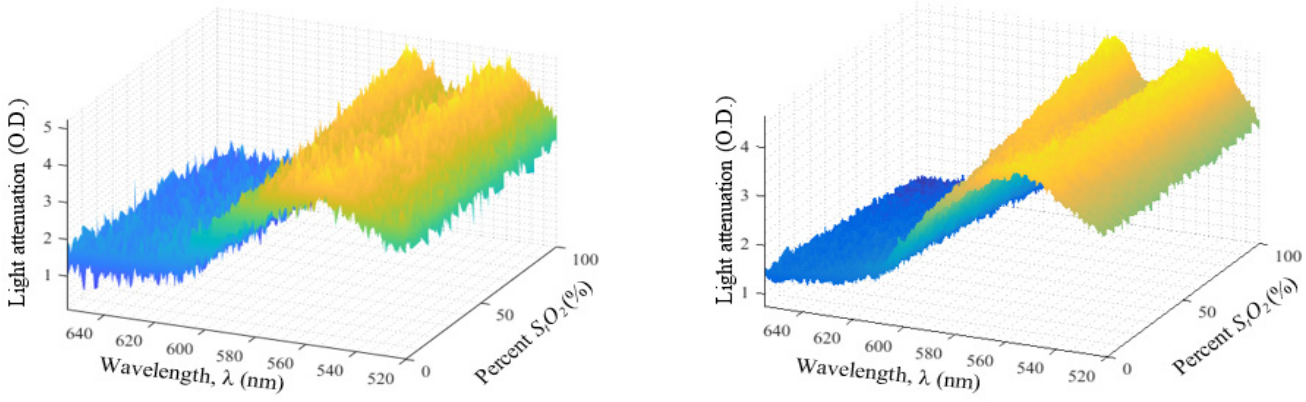


Figure 3. Three-dimensional surface plots of light attenuation values for SNR of (left) 20 and (right) 30 dB.

650 nm at a step resolution of 1 nm), this produces a poor network and association on the relationship between input and target values. The prediction errors can be minimized by considering features that are of utmost significance, such as those within the presence of troughs and crest across wavelengths 540 to 570 nm. An example of previous works from this laboratory in [22] extracted and considered the important features (i.e. magnitude and wavelength of the hemoglobin peaks) in the prediction work, hence minimizing errors from overfitting.

Meanwhile the KNN technique is primarily based on the location (i.e. magnitude) of these attenuation values, large difference in the values with changes in S_tO_2 would be desirable to yield better accuracy result. For this reason, relatively consistent errors were observed in Table 1. We hypothesized that this technique could be successfully applied to the classification of skin pigmentations.

It must be mentioned that the use of two-layered homogeneous medium is insufficient to represent human skin tissues since there are differences in optical properties of heterogeneous sub-layers in skin as widely reported in earlier works [27, 28]. Nonetheless the findings from this study would be useful when measurement data are readily available.

5. Conclusion

This work investigated the performance of different AI techniques in the prediction of medium S_tO_2 using MC technique. The significant findings of this study include the validation of the importance of hemoglobin signatures in the prediction of the required S_tO_2 value. The notably better accuracy in the values predicted using PLS regression technique is owing to its ability to distinguish the hemoglobin peaks using latent factor of 3. Meanwhile the shortcoming of KNN technique in the prediction work is largely contributed by the high similarity in the range and distribution of light attenuation values making the classification process complex and difficult. In addition, the large spectral information used in ANN technique resulted in over-fitting problem and produced comparatively larger errors. This work concluded that PLS regression performed well in the prediction of S_tO_2, and this technique may also be suitably used to select the distinctive signatures of spectral prior to the development of a prediction model using other AI technique.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to acknowledge Research Management Center (RMC) Universiti Tun Hussein Onn Malaysia for funding this study under research fund Vot E15501.

References

[1] R. M. Siao, M. J. So, M. H. Gomez, "Pulse oximetry as a screening test for hemodynamically significant lower extremity pheripheral artery disease in adults with type 2 diabetes mellitus" Journal of the ASEAN Federation of Endocrine Societies, **33**(2), 1-20, 2018. https://doi.org/10.15605/jafes.033.02.04

[2] N. T. Clancy, S. Saso, D. Stoyanov, V. Sauvage, D. J. Corless, M. Boyd, D. E. Noakes, M. Y. Thum, S. G. Maghami, J. R. Smith, D. S. Elson, "Multispectral imaging of organ viability during uterine transplantation surgery in rabbits and sheep" Journal of Biomedical Optics, 21(10), 1060061-7, 2016. https://doi.org/10.1117/1.JBO.21.10.106006

[3] S. Akter, S. Kawauchi, S. Sato, S. Aosasa, J. Yamamoto, I. Nishidate, "In vivo imaging of hepatic hemodynamics and light scattering property during ischemia-reperfusion in rats based on spectrocolorimetry" Biomedical Optics Express, 8(2), 972-992, 2017. https://doi.org/10.1364/BOE.8.000974

[4] C. Carlile, C. E. Wade, M. S. Baraniuk, J. B. Holcomb, L. J. Moore, "Evaluation of S_tO_2 tissue perfusion monitoring as a tool to predict the need for lifesaving interventions in trauma patients" The American Journal of Surgery, 210(6), 1070-5, 2015. https://10.1016/j.amjsurg.2015.08.006

[5] R. Rutkowski, M. Schuster, J. Unger, C. Seebauer, H. R. Metelmann, Th.v. Woedtke, K. D. Weltmann, G. Daeschlein, "Hyperspectral imaging for in vivo monitoring of cold atmospheric plasma effects on microcirculation in treatment of head and neck cancer and wound healing" Clinical Plasma Medicine, 7(8) 52-57, 2017. https://doi.org/10.1016/j.cpme.2017.09.002

[6] R. Re, I. Pirovano, D. Contini, L. Spinelli, A. Torricelli, "Time domain near infrared spectroscopy device for monitoring muscle oxidative metabolism: Custom probe and in vivo applications" Sensors, 18(1), 264, 2018. https://doi.org/10.3390/s18010264

[7] S. Zhang, S. Gnyawali, J. Huang, W. Ren, G. Gordillo, C. K. Sen, R. Xu, "Multimodal imaging of cutaneous wound tissue" Journal of Biomedical Optics, 20(1), 016016, 2015. https://doi.org/10.1117/1.JBO.20.1.016016

[8] A. Y. Mousa, J. L. Ballard, Transcutaneous Oxygen Tension: Principles and Applications, Springer Link, 2017

[9] M. B. Vestergaard, U. Lindberg, N. J. Aachmann-Andersen, K. Lisbjerg, S. J. Christensen, I. Law, P. Rasmussen, N. V. Olsen, H. B. W. Larsson, "Acute hypoxia increases the cerebral metabolic rate- a magnetic resonance imaging study" Journal of Cerebral Blood Flow & Metabolism, 36(6), 1046-1058, 2016. https:// 10.1177/0271678X15606460

[10] Z. Deng, C. Li, "Noninvasively measuring oxygen saturation of human fnger-joint vessels by multi-transducer functional photoacoustic tomography" Journal of Biomedical Optics, 21(6), 0610091-5, 2016. doi: 10.1117/1.JBO.21.6.061009

[11] M. Barberio, F. Longo, C. Fiorillo, B. Seeliger, P. Mascagni, V. Agnus, V. Lindner, B. Geny, A.L. Charles, I. Gockel, M. Worreth, A. Saadi, J. Marescaux, M. Diana, "Hyperspectral enhanced reality a physiology-based surgical guidance tool" Surgical Endoscopy, 1-9, 2019. https://doi.org/10.1007/s00464-019-06959-9

[12] Q. Wang, Q. Li, M. Zhou, Z. Sun, H. Liu, Y. Wang, "A hyperspectral vessel image registration method for blood oxygenation mapping" PLOS ONE, 2017. https://doi.org/10.1371/journal.pone.0178499

[13] W. Feng, R. Shi, C. Zhang, T. Yu, D. Zhu, "Lookup table based inverse model for mapping oxygen concentration of cutaneous microvessels using hyperspectral imaging" Optics Express, 25(4), 3481-3495, 2017. doi:10.1364/oe.25.003481

[14] A. Vogel, V. V. Chernomordik, J. D. Riley, M. Hassan, F. Amyot, B. Dasgeb, S. G. Demos, R. Pursley, R. Little, R. Yarchoan, Y. Tao, A. H. Gandjbakhche, "Using noninvasive multispectral imaging to quantitatively assess tissue vasculature" Journal of Biomedical Optics, 12(5), 051604 -,2007. https://doi.org/10.1117/1.2801718

[15] S. P. Philimon, A. K. C. Huong, X. T. I. Ngu, "Investigation of spatial resolution dependent variability in transcutaneous oxygen saturation using point spectroscopy system" in 2017 IOP Conf. Ser.: Mater. Sci. Eng. 226, 012122, 2017.

[16] S. P. Philimon, A. K. C. Huong, X. T. I. Ngu, "An alternative wavelength range for noninvasive assessment of wound tissue oxygenation status" International Journal of Engineering and Technology, 7 (4.26), 73-33, 2018. DOI: 10.14419/ijet.v7i4.26.22141.

[17] R. Liu, B. Yang, E. Zio, X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review" Mechanical Systems and Signal Processing, 108, 33- 47, 2018. https://doi.org/10.1016/j.ymssp.2018.02.016

[18] M. I. Jordan, T. M. Mitchell, "Machine leaning: Trends, perpecrtives, and prospects" Science, 349 (6245), 255 – 260, 2015. DOI: 10.1126/science.aaa8415

[19] H. Akbari, L. V. Halig, H. Zhang, D. Wang, Z. G. Chen, B. Fei, "Detection of cancer metastasis using a novel macroscopic hyperspectral method" in Proceedings of SPIE--the International Society for Optical Engineering, 8317, 831711. doi:10.1117/12.912026

[20] V. A. Convertino, S. L. Moulton, G. Z. Grudic, C. A. Rickards, C. Hinojosa-Laborde, R. T. Gerhardt, L. H. Blackbourne, K. L. Ryan, "Use of advanced machine learning techniques for noninvasive monitoring of hemorrhage" Journal of Trauma, 71(1), 25-32, 2011. doi: 10.1097/TA.0b013e3182211601

[21] R. Fusco, M. Sansone, S. Filice, G. Carone, D. M. Amato, C. Sansone, A. Petrillo, "Pattern recognition approaches for breast cancer DCE-MRI classification: A systemic review", Journal of Medical and Biological Engineering, 36(4), 449-459, 2016. DOI: 10.1007/s40846-016-0163-7

[22] A. Huong, X. Ngu, "Neural network approach for rapid prediction of transcutaneous oxygen saturation", in 2019 IEEE 9th Symposium on Computer Applications & Industrial Electronics (ISCAIE), Malaysia, 2019. DOI: 10.1109/ISCAIE.2019.8743751

[23] H. Y. Ma, Z. Zhou, S. Wu, Y. L. Wan, P. H. Tsui, "A computer aided diagnosis scheme for detection of fatty liver in vivo based on ultrasound kurtosis imaging" Journal of Medical Systems, 40(33), 1-9, 2016. DOI: 10.1007/s10916-015-0395-z

[24] A. Doronin, I. Meglinski, "The application of a unified Monte Carlo model in the training of artificial neural networks for the purpose of real-time in-vivo sensing of tissue optical properties" in 2019 SPIE Defense and Commercial Sensing, Baltimore, Maryland, United States, 2019. https://doi.org/10.1117/12.2518566

[25] A. Roggan, M. Friebel, K. Dorschel, A. Hahn, G. Muller, "Optical properties of circulating human blood in the wavelength range 400 – 2500 nm", Journal of Biomedical Optics, 4(1), 36-46, 1999. https://doi.org/10.1117/1.429919

[26] P. Mitra, C. A. Murthy, S. K. Pal, "Density-based multiscale data condensation" IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(6), 734-747, 2002.

[27] Y. Masuda, Y. Ogura, Y. Inagaki, T. Yasui, Y. Aizu, "Analysis of the influence of collagen fibres in the dermis on skin optical reflectance by Monte Carlo simulation in a nine-layered skin model", Skin Res Technol. 24(2), 1-8. 2017. https://doi.org/10.1111/srt.12421

[28] M. N. S. Yusiff, M. S. Jaafar, "Effect of multidesign skin model and characteristic on Monte Carlo simulation of light skin diffuse reflectance spectra", Jurnal Teknologi, 78(9), 135-142, 2016. DOI: 10.11113/jt.v78.7605

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 338-343 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Mutual Reduction in the Coupling of the MIMO Antenna Network Applied to the Broadband Transmission

Chafai Abdelhamid[1], Hedi Sakli*,[1,2]

[1]*MACS Research Laboratory LR16ES22, National Engineering School of Gabes, Gabes University, 6029 Gabes, Tunisia.*

[2]*EITA Consulting, 5 Rue du Chant des oiseaux, 78360 Montesson, France.*

ARTICLE INFO

Article history:
Received: 13 December, 2020
Accepted: 18 March, 2020
Online: 26 March, 2020

Keywords:
UWB antenna
MIMO antenna
SRRs metamaterial

ABSTRACT

In this article, a new form of UWB (ultra-wide-band) antenna operating to the desired specifications, obtained from a base antenna to which some modifications are made. The proximity of the antennas causes a mutual coupling phenomenon thus generating an apparent modification of their characteristics. It is therefore crucial to have the minimum level of insulation with a small possible separation distance between the antennas to ensure efficient operation of our multiple input and multiple output (MIMO) system. The separation spacing between the antennas is close to 0.15λ0 (λ0 is the wavelength in free space), where the coupling can be very weak. To overcome this weakness, two efficient methods are applied to minimize mutual coupling in a four-element MIMO antenna, such as the use of the neutralization line and the use of split ring resonator metamaterial (SRR). The objective of our paper is, on the one hand the design of UWB antenna, and on the other hand to minimize the coupling between four UWB-MIMO antennas using the two methods of isolation mentioned above. A coupling minimization of more than 15 dB is obtained using the neutralization line and more than 20 dB with the use of metamaterial while maintaining the operating band and radiation specifications of the MIMO antennas.

1. Introduction

In recent years and today, the MIMO antenna has demonstrated its ability to improve the efficiency and capacity of the satellite communication system [1]. Despite the advantages of the MIMO antenna in modern technology, mutual coupling is a delicate challenge, given the small distance between the elements [2,3]. The effects of the interaction between all components of the MIMO antenna array have the effect of reducing the operation of the system by increasing the correlation coefficient and the channel loss. This reduces the antenna gain and distortion of the antenna pattern. Radiation [3-5], and consequently reduces the efficiency of the MIMO antenna. The influence of the near fields performed by the neighboring antenna with the presence of the surface waves has a great influence on the radiating elements in the emission-reception system and more precisely on the coupling between antennas [1,6]. In cases where the antennas are placed on dielectric substrates with low permittivity [3], the fields of the antennas are very close, a very strong coupling these produce. The problems related to antenna linkages have increased since the end of the 20th century. Many research studies have been conducted and are continuing to improve the characteristic radiation efficiency of antennas. Metamaterial are a promising research topic in various fields, in particular electromagnetism [7], whether for circuit applications (filters, phase shifters, etc.) or for radiation applications (antennas, diffraction, stealth). These are elaborate materials, consisting of periodic elements of small size (relative to the wavelength) and providing, in specific frequency bands, particular properties different from those of natural materials [8,9]. Potential benefits of these structures include, for example, miniaturization of antennas, expansion of their bandwidth or reduction of inter-element coupling within a network. To take into account the constraints inherent in the need for discretion and to respond to the current trend of low thickness solutions, we favored a planar technology for the design of this new antennal solution. The delicate compromise between bandwidth, dimensions and electrical performance leads us to propose a new radiating print. In this article, we propose a very wide bandwidth rectangular patch antenna on a FR4 substrat, with insertion of the notches at the ground plane and shape modifications for the radiating element. Compared with conventional microstrip antennas, the

*Corresponding Author : Hedi Sakli, hedi.s@eitaconsulting.fr

bandwidth is around 15 GHz for a reflection coefficient of -10 dB. A physical structure is proposed to minimize the mutual coupling between the elements constituting the MIMO system with four patches very close to each other. Two isolation techniques are studied and compared in this article. Starting with the use of an isolation technique called DGS (Defected Ground Structure) is to introduce changes to the structure of the ground plane to change the current distribution. In MIMO systems, the effect of the notch filter is applied in order to reduce the mutual coupling between the radiating antennas. The application of the DGS method is to introduce slits on the ground plane [10-12] The DGS solution is very easy to implement since its operation depends on the resonance frequency rather than the types of antennas, so a combination of this technique with the technique called "neutralization technique" (NL) used to decouple the two UWB antennas. MIMO placed side by side and involves inserting a line between the elements in a well-defined position or short-circuit when the antenna are positioned very closely on a small mass plane. Noting that the electromagnetic coupling presented between two patchesnearby is mainly of the capacitive type [12,13]. Indeed, by introducing an inductance between the two patches, a rejection filter behavior allows decoupling [14,15]. A decoupling of the order of 14 dB is obtained in the operating band [3.2 - 17.7] GHz. The second solution is to use a structure based on metal rings (Split Ring Resonators or SRR) with a permeability of negative value. Four SRRs in series are located along the distance between two adjacent antennas and two opposite antennas forming the symbol (+) based on the upper layer of the FR4 substrate to separate the antennas from each other to ensure good isolation. Mutual coupling between the elements of the MIMO antenna systems can be greatly reduced, with isolation greater than 10 dB over the majority of the antennas operating band.

In this paper, the design of UWB antenna is presented. The coupling between four UWB-MIMO antennas using the two methods of isolation (NL and metamaterial) is minimized. The gain of isolation betwen antennas more than 15 dB using the neutralization line and more than 20 dB with the use of metamaterial are obtained while maintaining the operating band and radiation specifications of the MIMO antennas.

2. Antenna characterization

We start at the beginning of our design with a square-shaped printed antenna powered by microstrip line. The choice of the power line is based on the fact that this line must have a characteristic impedance of 50 ohms, adapted over a very wide frequency band. The geometry of this antenna consists of a rectangular patch printed on the top face of a FR4 type substrate. The ground plane is a partial plane printed on the bottom surface of the substrate. The reflection coefficient modulus of the initial antenna is greater than -10 dB for frequencies in the band [3 - 11] GHz. But in the band [11 - 16.7] GHz the modulus of reflection is less than -10 dB (see figure 1 curve a). This antenna is therefore not well suited over the entire frequency band studied (X, Ku and Ka bands). According to the influence of the modifications made on the base antenna, and in order to improve its characteristics and mainly its adaptation as well as obtaining an ultra-wide band operation The design is made by introducing some modifications on the structure of the base antenna, in particular on the shape of the radiating element and by adding slots on the patch and on the ground plane; we present in this part our complete antenna with all the modifications made. Figure 2 presents the geometric shape as well as the dimensions of our antenna presented. The different characteristics of the antenna will be presented and commented. Figure 1 presents the geometry as well as the dimensions of the proposed antenna printed on a substrate FR4 of relative permittivity εr = 4.4, loss tangent tan δ = 0.02 and thickness equal to 1.6 mm. The parameters associated with the structure of the antenna studied are shown in Table 1.

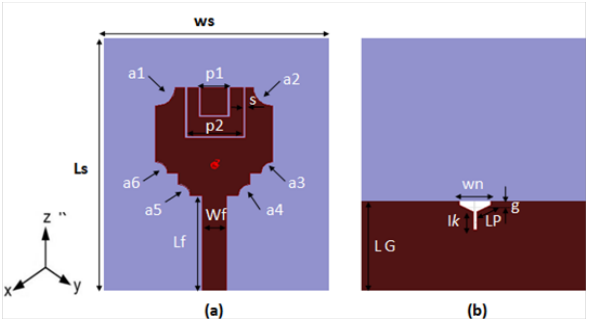


Figure 1: Structure of the patch antenna studied: (a) In front view (b) bottom view.

Table 1: Geometric parameters of the studied antenna.

Parameters	Values (mm)	Parameters	Values (mm)
W_S	25.52	L_S	31
W_{Pa}	14.52	L_{Pa}	14.52
W_F	2.85	L_F	13.52
L_G	12.62	H_{Su}	1.6
a1	1.45	g	0.36
a_2	1.5	l_l	7.5
a_3	1.1	L_p	1.62
a_4	1.1	wn	3.1
a_5	1.1	L_k	0.75
a_6	1.1	L_p	2.35
p_1	10		
p_2	4		

This antenna has a reflection coefficient of -10 dB over the entire frequency band [3.1 - 16.8] GHz (Figure 2 curve b).

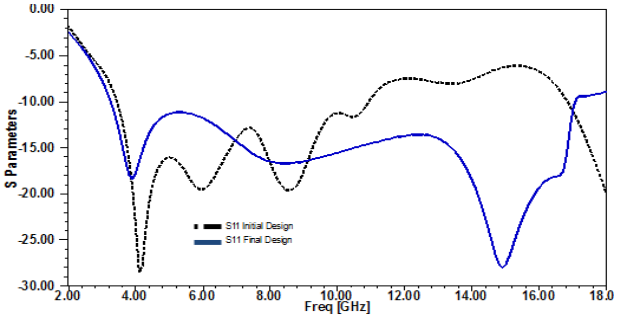


Figure 2: Reflexion coefficient S11 of the UWB antenna: (a) Initial design, (b) Final design.

The influence of modification is clearly visible in the improvement of the adaptation.

3. NL Decoupling structure design

At the beginning a research was done to show the effect of the neutralization line on the general radiating element and more

precisely on the insulation between the elements that constituted our system. The basic idea of this method was to compensate for the complex electromagnetic coupling existing between the neighboring antennas by opposite coupling reaction. As the coupling between the antennas is capacitive, the line mainly makes it possible to provide a reaction or an inductive effect to oppose the capacitive effect of the coupling between antennas [16]. This technique has been used to guarantee improved isolation of several multi-antenna systems [17]. Moreover, the neutralization line supports strong currents so that the direction of current is directed towards the antenna itself to radiate and not towards the power port of the second antenna. In addition, it is possible to cancel the coupling of all the paths [16]. Then this method is the one selected to improve the isolation of the four patch antennas of our multi-antenna system. The insertion of the line is done in the area where the amplitude of the currents is the highest, which corresponds to the zones of low impedances. A simple four-element structure has been used. The width and length of the microstrip antenna are the same. The separation distance between the four elements is 0.15λ0. The MIMO antenna structure is based on a FR4 type substrate 60 x 48 mm in size with a permittivity (εr) of 4.4 mm and a thickness of 1.6 mm. The proposed antenna scheme is shown in figure 3. NL was placed between items like shown in Figure 4.

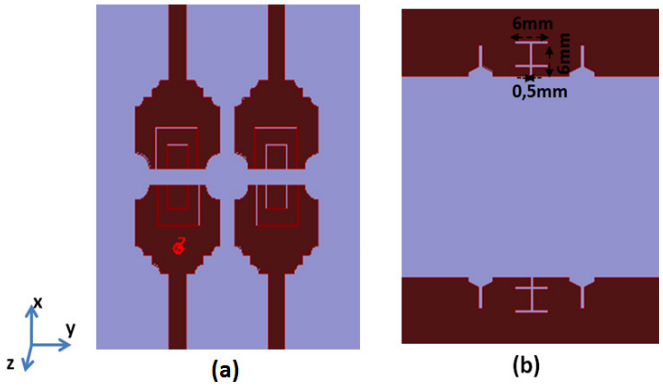


Figure 3: MIMO studied with four elements without NL.

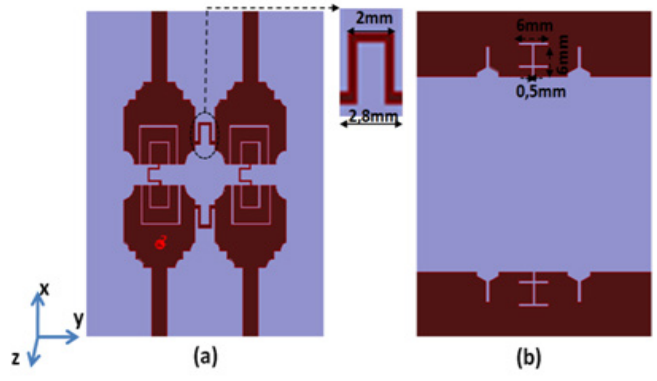


Figure 4: MIMO studied with four elements: with NL

3.1. Reflexion coefficient (Sii) Analysis

Several simulations have been carried out to find the optimal position of the neutralization line to maintain the bandwidth in adaptation. The simulation result has shown a small assignment on the parameters S11, S22, S33 and S44, which are in coincidence. The scattering parameters simulated are shown in figure 5. By comparing the two simulation results without and with NL, it is noted that the parameters of the Sii reflection coefficients (S11, S22, S33 and S44 are the same because of the symmetry of the structure) are not strongly modified, the MIMO antenna covered a very wide bandwidth of - 10 dB.

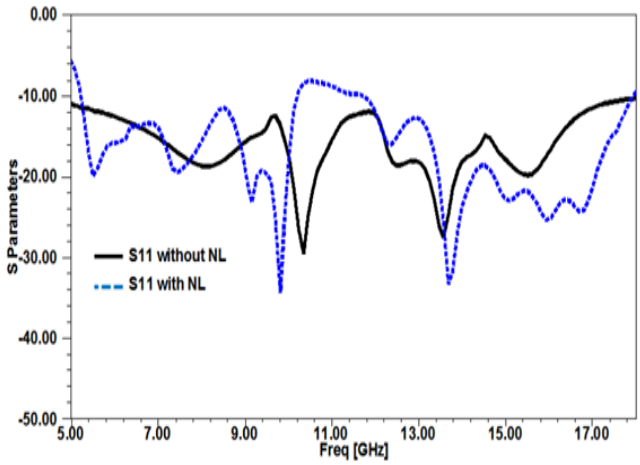


Figure 5: Reflexion coefficient S11 of the UWB antenna without and with NL

3.2. Isolation Parameters Analysis

The transmission parameters Sij from antenna j to antenna i (S21, S31 and S41) reveal and show that the adjacent antennas are strongly coupled with a decoupling method that is easy to design and execute compared to that presented in several articles. The parameter level is decreased for our current MIMO system. According to the results obtained by simulation, it has been found that the influence of an NL makes it possible to find an acceptable decoupling level for a network of four antennas. This helps us to create a model similar to the one created by simulation. Compared to the MIMO-UWB antenna without NL application, our studied MIMO antenna not only preserves the overall performance for each element, but also further minimizes the coupling of 21 dB at 15 GHz and 15 dB at 8 GHz and presents an isolation more than 15 dB, (see figures 6, 7 and 8).

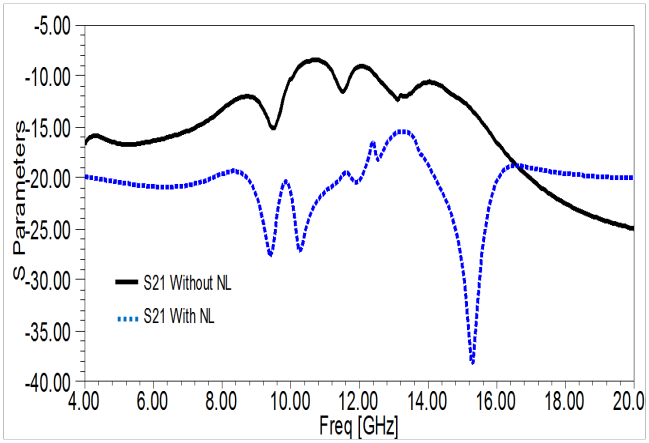


Figure 6: Parameters S21 simulated of MIMO antennas (adjacent) without and with NL

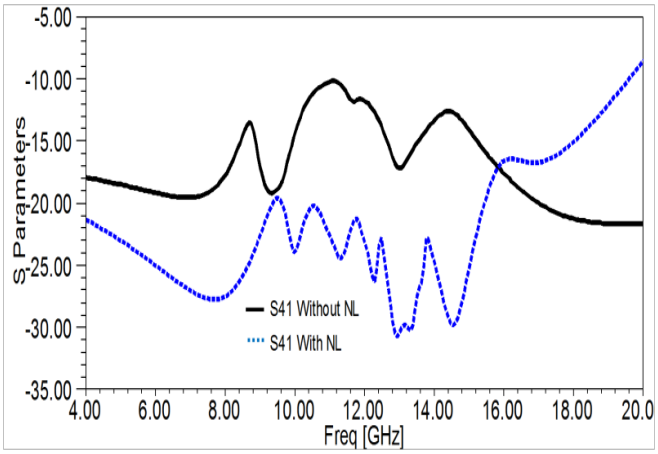


Figure 7: Parameters S41 simulated of MIMO antennas (opposite) without and with NL

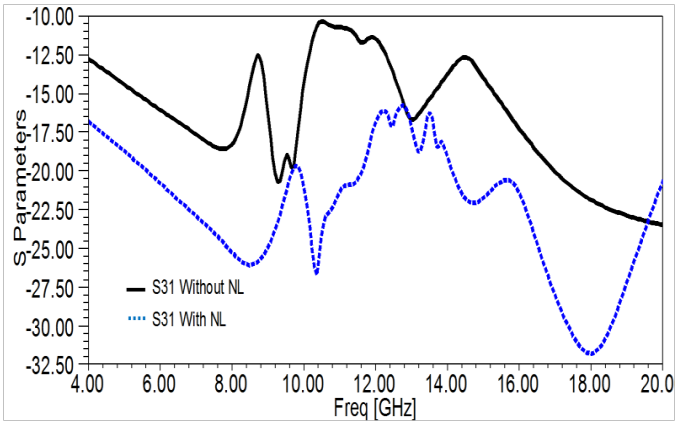


Figure 8: Parameters S31 simulated of MIMO aantennas (symmetrical) without and with NL

4. Structure and decoupling of metamaterials

Research has been carried out to ensure and obtain an electromagnetic mulberry allows to isolate the flux and the interference to the radiation of proximity of adjacent antennas. Subsequently, a proposed method based on the application of the metamaterial was presented which makes it possible to minimize the coupling levels [17]. So at the beginning a design of a SRR presented to ensure the operation around the band of frequencies mentioned initially and are suitable the impact on the resonance of the components of the MIMO antenna. The applied SRR unit cell has been dimensioned for band operation. The substrate used for the simulation is the FR4 which has a relative permittivity of 4.4, tangential losses of the order of 0.002 and a thickness of 1.6 mm. The external side of the square is equal to 2.4 mm. The width of the track in copper is 0.3 mm. The width of the cut of the rings is 0.34 mm (figure 9).

With the aim of guaranteeing isolation over our entire frequency band, we are doing a study on the size of the SRR aperture in order to see the influence of this variation on the resonance frequency of SRR. According to the simulation results mentioned in the figure 10 and figure 11, we notice well that we have a relationship of proportionality between the increasing of the opening for the two rings of our SRR and the resonant frequency that evolves with the opening.

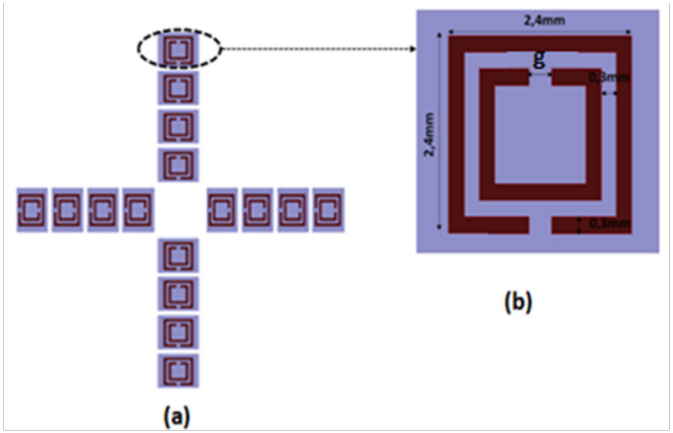


Figure 9: Geometric of studded split ring resonators: (a) 4×4 unit cells in front view and (b) Unit cell

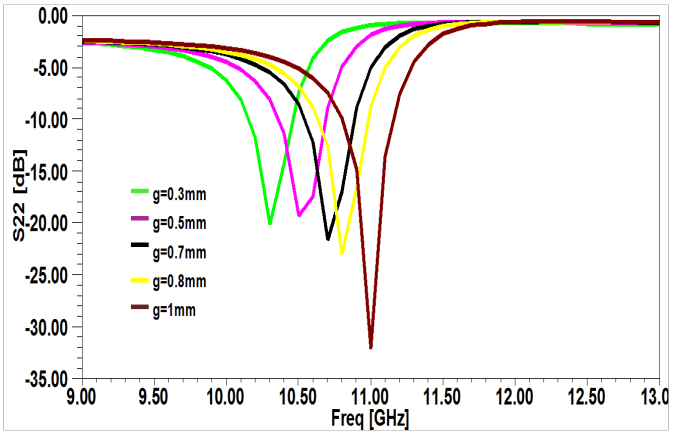


Figure 10: Parameter of S22 for several aperture values in the SRR

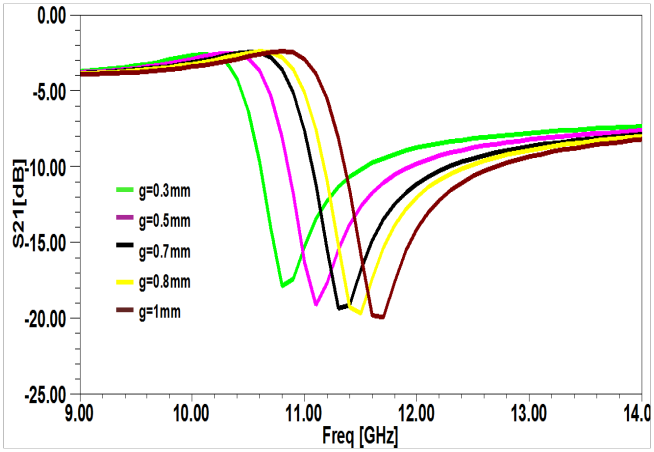


Figure 11: S21 parameters for several aperture values in the SRR

The structures are placed between the antennas at an optimized separation distance at 0.15λ0 and are composed of 4 groups of SRRs over the substrate. Each group has four integrated SRR units between the antennas for good isolation between the antenna elements. This configuration is illustrated in Figure 12. In this part, we are going to study the influence of an SRR cell network between the two adjacent antennas. We have looked for presenting an electromagnetic blackberry in order to isolate radiation fluxes and interferences from the proximity of adjacent antennas. So a proposed solution based on the use of the

metamaterial has been presented which allows stopping or minimizing the propagation of surface waves, thus reducing the mutual coupling in a MIMO antenna.

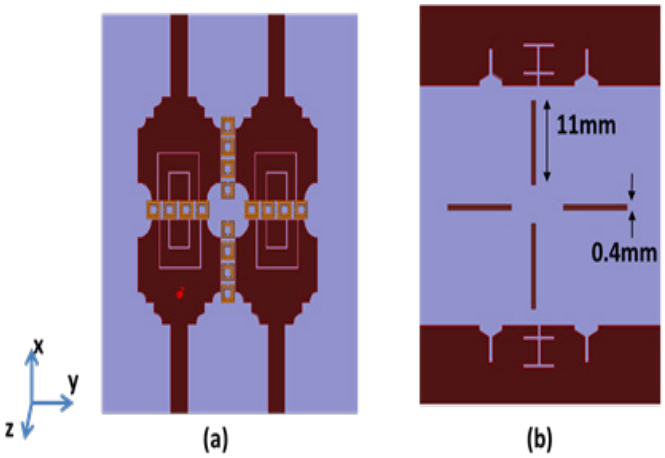


Figure 12: UWB MIMO antenna associated with SRR cells in front view and bottom view.

4.1. Reflexion coefficient (S11) Analysis

We note that the simulated structure has a stability of the S11 reflection coefficients of four MIMO antennas with SRRs over the entire desired band. Figure 13 indicates the parameters S11 of the proposed antenna. The influence of the material is clearly visible in the improvement of the adaptation of the antenna. We can conclude that the structure with the application of the SRR contributed to the increase of the MIMO antenna performance in terms of reflection coefficient and consequently allowed us to obtain a better ULB structure.

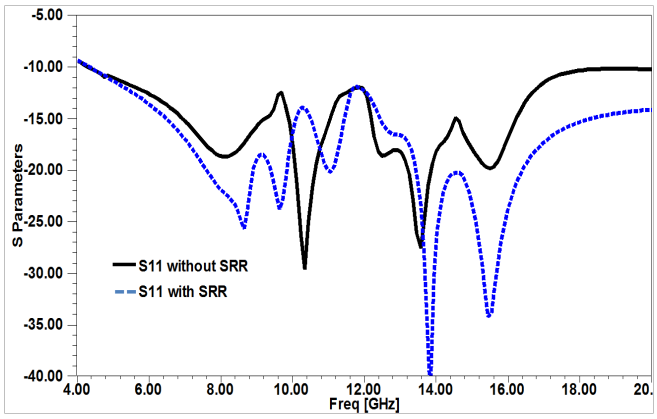


Figure 13: Reflexion coefficient S11 of the UWB antenna with SRR.

4.2. Isolation Parameters (S_{i1}) Analysis

Inter-element electromagnetic coupling presents a difficulty between the components during start-up. Here, a chain of separation between the elements is an assembly of two periodic networks consisting of metallic line and a resonator network makes it possible to have a negative permeability whereas the line network is characterized by a negative permittivity. The assembly of these two networks contributes to having a negative permittivity and permeability at the same time. The SRR chain at the highest point of the substrate was used to ensure isolation improvement. As previously reported, the mutual coupling is more important in the case where the radiating elements are very close. For this reason, we are interested in this approach to the analysis of the mutual coupling between each two adjacent, opposite and symmetrical antennas. In fact, as shown in figures 14, 15 and 16, the initial mutual coupling measured in terms of transmission coefficient at the frequency of 10.7 GHz was -7.2 dB, -10.1 dB and -11.6 dB between antenna 1 and 2, 1 and 3, 1 and 4 respectively. After the use of SRR, the isolation becomes more improved to obtain the value of -24.3 dB, -22.2 dB and -27.6 dB between antenna 1 and 2, 1 and 3, 1 and 4 respectively. Thus, the condition of having a mutual coupling lower than -17 dB imposed to maintain the performance of a MIMO system.

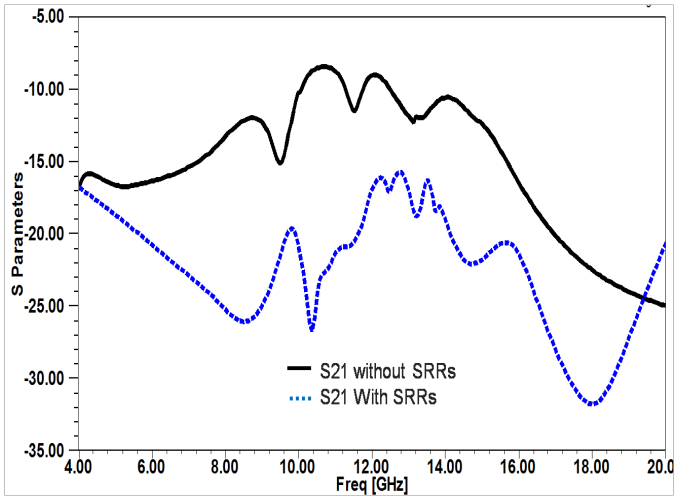


Figure 14: Parameters S21 simulated of MIMO antennas (adjacent) without and with metamaterial.

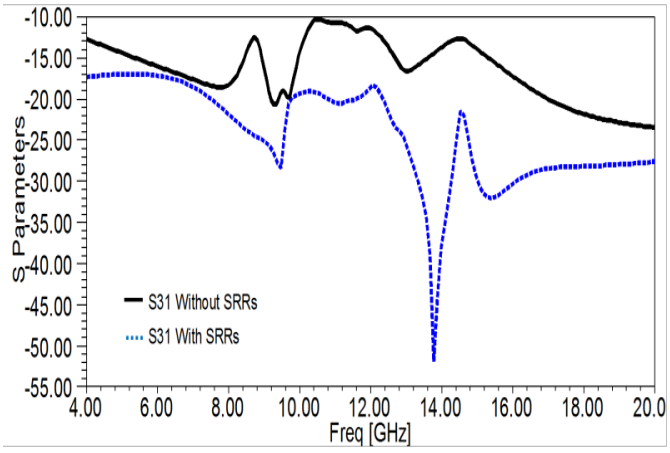


Figure 15: Parameters S31 simulated of MIMO antennas (symmetrical) without and with metamaterial.

The observation on the figures 14, 15 and 16 showing the transmission coefficients indicates better quality with reduced power losses across the desired band. An isolation protection gain of more than 10 dB over the entire band is achieved. This is explained by the isolation performance between the antennas with this predicted structure. This separation between the antennas ensuring the efficiency of the transmission system constituting our space links.

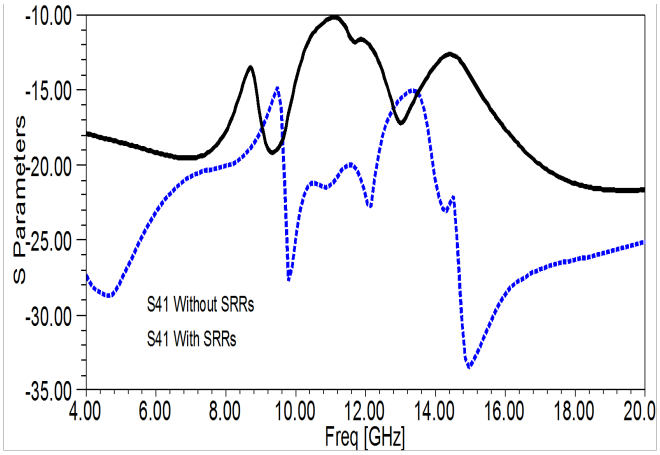


Figure. 16: Parameters S41 simulated of MIMO antennas (opposite) without and with metamaterial.

5. Conclusion

In this article, we presented a new structure of a MIMO antenna array that belongs to the UWB ultra-wideband UWB antennas categories. We presented several approaches to improve the performance of MIMO systems. The simulation relating to a MIMO antenna has been completely described, have shown great efficiency for the improvement of the isolation of a system consisting of four UWB antennas originally presented. The results of electromagnetic simulation have shown the gain provided by these techniques for increasing the isolation between antennas in the band of interest. A solution that seems to be relevant is obtained using the neutralization line technique. For this solution, the insulation in the UWB band is greater than 18 dB. Another solution made it possible to highlight the role of the split-ring resonator, in the design of UWB microband antennas with the insertion of periodic SRR structures. We were able to increase the isolation without losing the UWB characteristics. These results show that the proposed structures contribute to meeting the requirements of modern communication systems dedicated to 5G.

References

[1] M. A. Vázquez, *et al*, "Precoding in multibeam satellite communications: Present and future challenges," IEEE Wireless Commun., vol. 23, no. 6, pp. 88–95, Dec. 2016.

[2] B. H. G. Gao, L. He, S. Wang, and C. Yang, "Investigation of A Reconfigurable Dual Notched UWB Antenna by Conceptual Circuit Model and Time-Domain Characteristics," Microwave and Optical Technology Letters, vol. 59, pp. 1326– 1332, 2017.

[3] A. Iqbal, O. A. Saraereh, A. W. and S. Bashir, "Mutual Coupling Reduction Using F-Shaped Stubs in UWB-MIMO Antenna", IEEE Access vol. 6, pp. 2755–2759, 2018.

[4] I. Szini, A. Tatomirescu, and G. F. Pedersen, "On Small Terminal MIMO Antennas, Harmonizing Characteristic Modeswith Ground Plane Geometry," IEEE Antenna Propag. Trans. On, vol. 63, no. 4, pp. 1487 - 1497, 2015.

[5] B. H. G. Gao, L. He, S. Wang, and C. Yang, "Investigation of A Reconfigurable Dual Notched UWB Antenna by Conceptual Circuit Model and Time-Domain Characteristics," Microwave and Optical Technology Letters, vol. 59, pp. 1326–1332, 2017.

[6] L. Malviya, R. K. Panigrahi and M. V. Kartikeyan., "A 2x2 Dual-Band MIMO Antenna with Polarization Diversity for Wireless Applications," Progress In Electromagnetics Research C, vol. 61, pp. 91-103, 2016.

[7] A. Dadgarpour, M. S. Sorkherizi and A. A. Kishk, "Wideband, Low loss Magneto Electronic Dipole Antenna for 5G Wireless Network with Gain Enhancement Using Meta Lens and Gap Waveguide Technology Feeding," IEEE Transactions on Antennas and Propagation, vol. 64, no. 12, pp. 5094 - 5101, 2016.

[8] M. S. Sharawi, S.K. Podilchak, M. T. Hussain and Y. M. M. Antar, "Dielectric Resonator Based MIMO Antenna System Enabling Millimeter-Wave Mobile Devices," IET Microwaves, Antennas& Propagation, vol. 11, no. 2, pp. 287 – 293, Jan. 2017.

[9] G. K. S. Pandey, H. S. Bharti, P. K. Meshram, M. Kumar, "Design and Analysis of Ψ-Shaped UWB Antenna with Dual Band Notched Characteristics," Wireless Personal Communications, vol. 89, pp. 79-92, 2016.

[10] R. B. Rani and S. K. Pandey, "A CPW-Fed Circular Patch Antenna Inspired by Reduced Ground Plane and CSRR Slot for Uwb Applications with Notch Band," Microwave and Optical Technology Letters, vol. 59, issue 4, pp. 745-749, 2017.

[11] G. K. Pandey, H. S. Singh, P. K. Bharti and M. K. Meshram, "Design and Analysis of Multiband Notched Pitcher-Shaped UWB Antenna" International Journal of RF and Microwave Computer-Aided Engineering,vol. 25, pp. 795–806, 2015.

[12] A. P. Singh, R. Khanna, and H. Singh, "UWB Antenna With Dual Notched Band For WIMAX and WLAN Applications," Microwave and Optical Technology Letters vol. 59, pp. 792– 797, 2017.

[13] T. Jiang, T. Jiao, and Y. Li, "Array mutual coupling reduction using L-loading e-shaped electromagnetic band gap structures," International Journal of Antennas and Propagation, vol. 2016, Article ID 6731014.

[14] S. Zhang and G. Pedersen, "Mutual coupling reduction for UWB MIMO antennas with a wideband neutralization line," IEEE Antennas and Wireless Propagation Letters, vol. 15, pp. 166-169, 2016.

[15] C. Abdelhamid, M. Daghari, H. Sakli and C. Hamrouni," A New UWB-MIMO Multi-Antennas With High Isolation For Satellite Communications," 15th International Wireless Communications & Mobile Computing Conference (IWCMC), Maroc, pp. 152-155, 24-28 Jun 2019. (978-1-5386-7747-6/19/$31.00 ©2019 IEEE).

[16] T. Jiang, T. Jiao, Y. Li, and W. Yu, "A low mutual coupling MIMO antenna using periodic multilayered electromagnetic band gap structures," Applied Computational Electromagnetics Society Journal, vol. 33, no. 3, pp. 305-311, 2018.

[17] K. Wei, J. Li, L. Wang, Z. Xing, and R. Xu, "Mutual coupling reduction by novel fractal defected ground structure band gap filter," IEEE Transactions on Antennas and Propagation, vol. 64, pp. 4328-4335, 2016.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 340-345 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Design of Efficient Convolutional Neural Module Based on An Improved Module

Daihui Li*, Chengxu Ma, Shangyou Zeng

Guangxi Normal University, College of Electronic Engineering, 541000, China

ARTICLE INFO

Article history:
Received: 05 November, 2019
Accepted: 19 January, 2020
Online: 07 February, 2020

Keywords:
Deep learning
Convolutional neural network
Feature extraction

ABSTRACT

In order to further improve the feature extraction performance of the convolutional neural networks, we focus on the selection and reorganization of key features and suspect that simple changes in the pooling layers can cause changes in the performance of neural networks. According to the conjecture, we design a funnel convolution module, which can filter out the key features and perform multi-scale convolution of key features. And we apply this module to the design of high performance small neural networks. The experiments are carried out on 101_food and caltech-256 benchmark datasets. Experiments show that the module has higher performance than classical module, and the small convolution networks based on the module has less parameters and more excellent performance.

1. Introduction

In 2012, deep learning shined on the ImageNet large scale visual recognition competition (ILSVRC), and more and more researchers began to notice deep learning. As an important branch of artificial intelligence, deep learning has been widely used in information processing in the fields of speech, image and text. Convolutional neural networks are the most classical network structures in the field of image processing, and they has excellent spatial feature learning abilities. The classical method to improve the convolution neural networks is to deepen and broaden the convolution neural networks, but it is easy to cause a large increase in the number of neurons. Large increases in the convolutional neural networks of parameters can lead to over-fitting, so the performance of neural networks will decline. Therefore, the model design of convolutional neural networks is particularly important. The excellent algorithm model can greatly improve the recognition accuracy and effectively reduce the computational complexity of the model, which determines the performance of almost all deep learning algorithms. However, the neural networks has so far been regarded as a black box device. In many cases, neural networks are not interpretative, and their design and improvement only rely on experience.

In this paper, we have improved the Reticulated Convolution Module (RCM) [1] and designed a small high performance convolutional neural network (HPCNet) with less parameters. Networks with fewer parameters are more easily ported to devices with limited transmission bandwidth and storage, such as cloud services, mobile platforms and smart robots. In addition, in the process of improved RCM, the pooling layer of the convolutional neural networks have attracted our attention. Simple changes in the use of different dimensional features lead to significant performance gaps in convolutional neural networks. This difference is instructive for the interpretability of convolutional neural networks and the design rules of convolutional neural networks. This paper mainly includes three parts of work: 1. We improved RCM and named it the second version of Reticulated Convolution Module (RCM-V2). 2. We designed a small convolutional neural network structure called backbone network (BBNet). And based on RCM-V2 and BBNet, a small convolutional neural network with high performance (HPCNet) was designed. 3. Finally, We evaluated RCMNet-V2 and HPCNet performance on two typical image classification datasets.

2. Related Work

Because of the birth of AlexNet [2], the design of convolutional neural networks has become an important research direction for scholars in the field of machine vision. The network structure of AlexNet has eight layers, including five convolution layers and three full connection layers. AlexNet adopts data enhancement methods and adds a dropout layer to prevent over-fitting. And it won the championship with a recognition accuracy of 10.9% ahead of its nearest rival. Since then, ILSVRC competitions had produced a large number of excellent network structures. VGGNet [3] explores the correlation between deep network and performance by continuously stacking small convolution kernels and pooling layers. GoogLeNet [4] transforms full connection and even general convolution into sparse connection to solve the defect,

* Daihui Li, Email: ldhcv@sina.com

but it keeps dropout. Two auxiliary softmax are added to the networks for forward conduction gradient to prevent the gradient from disappearing. GoogLeNet innovatively proposes the Inception structure which gathers the features of different scales together, expands the network width and enhances the expressive power of the networks. However, the deeper the network is, the more likely it is to cause degradation. In order to solve this problem, ResNet [5] came out. It introduced identity mapping into network structure and surpassed human in top-5 precision. ResNet is a milestone network, which makes it possible to train deeper networks. DenseNet [6] was born on CVPR2017 which established the dense connection between the front layers and the back layers. The connection of features between channels enables the network to realize high-density feature reuse, which achieves the performance beyond ResNet under the condition of less parameters and calculation cost.

With the increasing scale of the networks, resource asymmetry and other problems are caused in most scenarios. Therefore, compression and acceleration of algorithm models under the guarantee of network performance is a hot research topic in the field of network structure optimization. Lin M et al. [7] first proposed 1x1 convolution kernel in networks in order to greatly lessen network parameters. Iandola, F N et al. [8] proposed a lightweight network called SqeezeNet which got the AlexNet-level accuracy. Howard A G et al. [9] put forward the depthwise separable revolution in MobileNet. The depthwise separable revolution greatly reduces the parameters when the precision is slightly lower than the traditional convolution. However, the depthwise separable convolution reduces the information integration performance of the networks to a certain extent. Later, the idea of ResNext [10] was introduced into ShuffleNet [11] to clean the channels of convolution neural networks. Shufflenet strengthens the information circulation between channels, and improves the information expression ability of networks. The iconic NAS [12] promoted the spring tide of automatic machine learning (AutoML), which searched neural network structure according to reinforcement learning such as ENAs [13], DARTS [14] and NAO [15], etc. These models generated by automatic machine learning are not universal. It is also inconvenient to deploy an automated machine learning framework on some simple tasks, so designing an efficient small-convolution neural network is especially important. This paper aims to solve the problems of model design in simple machine vision tasks, and proposes an efficient micro convolution neural network called HPCNet. At the same time, some design inspirations of convolutional neural networks are obtained, which will provide inspiration for the design work of convolutional neural networks in the future.

3. Design of Small High Performance Convolutional Neural Network

This chapter first introduces the group convolution and the group convolution used in RCM. Then we introduce the design of RCM-V2 and its key features extraction process. Finally, the overall architecture of BBNet is introduced.

3.1. Group Convolution

Group convolution plays a crucial role in the design of RCM and RCM-V2. As shown in Figure 1, the group convolution groups the input features and then performs convolution operations within each group. When the size of the input feature maps is H*W*C and the output quantity of model is N, these feature maps are to be divided into G groups. After grouping convolution operation, each set of feature maps is C/G, and the output is N/G. It can be seen that the group convolution can consolidate information in small groups, which functions as feature isolation and cross interaction, and reduces the amount of parameters.

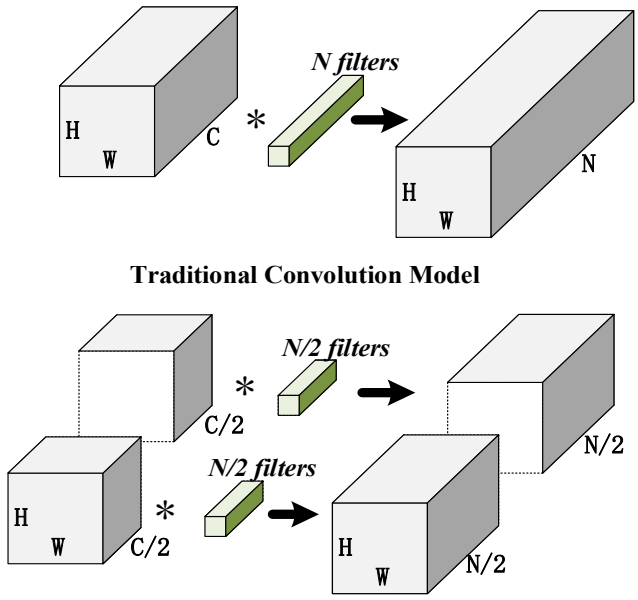


Figure 1. Group Convolution Model and Traditional Convolution Model

If the input of the convolution layer is composed of two different sets of features, and the number of feature maps in these two groups is not equal, great effect will be produced. As shown in Figure 2, A and B are two different sets of features, and the number of feature maps contained in A and B is different. A contains 128 feature maps and B contains 64 feature maps. Group convolution concatenates A and B, and then divides A and B into C and D. Both C and D contain 96 feature maps, and the feature maps in C are all derived from A. In D, 32 feature maps are derived from A, and 64 feature maps are derived from B. The convolution operation æ performs feature extraction on C, which is equivalent to feature selection and fusion on the first 96 convolution maps of A. The convolution operation £ performs feature extraction on D, which is equivalent to the feature fusion of the 32 feature maps of A and all the feature maps of B. As a result, F aggregates the features of A and B, while E merely extracts the features of A. Therefore, the features extracted by E and F will be significantly different.

3.2. Design of RCM-V2

The traditional convolution neural networks are formed by alternately stacking convolution layers and pooling layers. The desired effect is achieved based on the design of the kernel size of convolution layers and the design of different pooling layers. Single channel and single scale convolution settings tend to make the feature extraction of neural networks inadequate and constrain the performance of neural networks. In this paper, the overall structure of the mesh convolution module is funnel-shaped mesh.

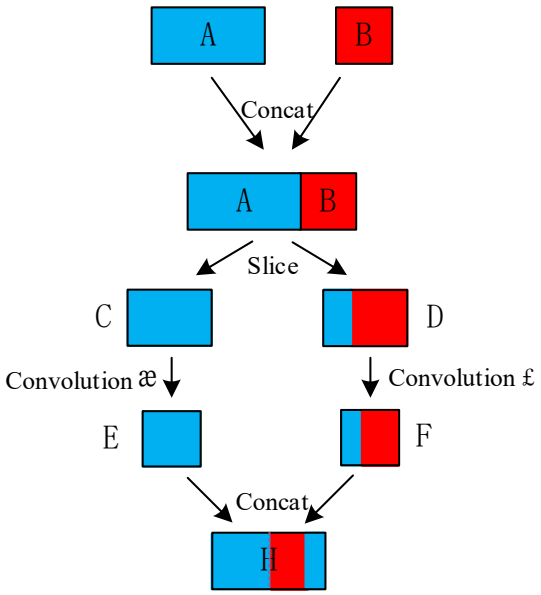


Figure 2. Group convolution of different feature sources.

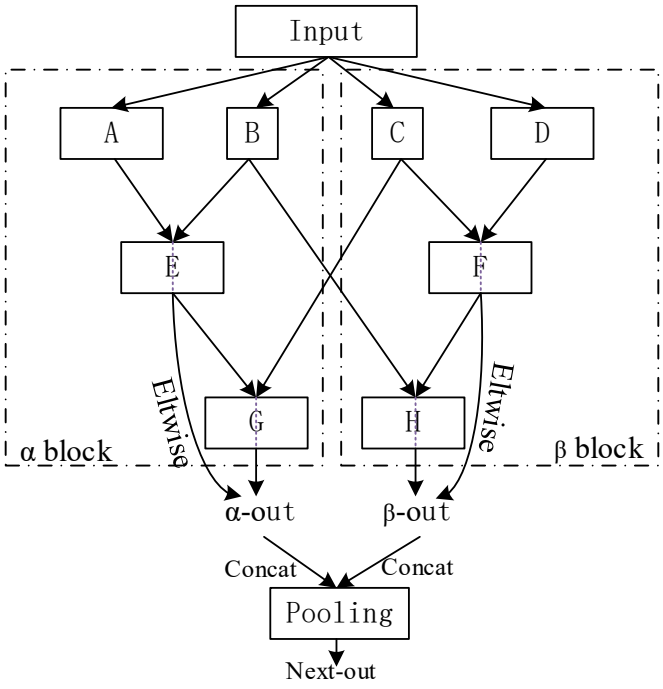


Figure 3. The structure of the RCM.

The overall structure of RCM is symmetrical, and the number of convolution kernels is unbalanced. RCMNet first integrates different features, and then combines them through cross structure. It not only realizes feature diversification, but also reuses a large number of important features, which makes convolution neural effectively enhance the ability of extracting key features. For miniature machine vision tasks, RCM can extract key features efficiently and ensure the diversity of features, so that convolution kernel can be used efficiently. Figure 3. shows the overall architecture of RCM. RCM is divided into α blocks and β blocks, and the key features are output of the bottom layer. In general, max-pooling preserves the main features. We focus on the pooling layers to further strengthen the feature selection ability of RCM.

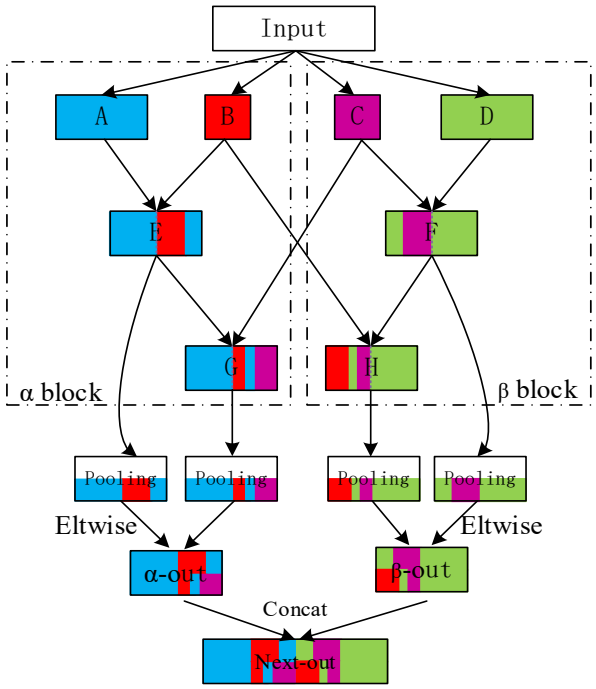


Figure 4. The structure of the RCM-V2. The number of groups for E, F, G and H layers is 2. In addition to the stitching method indicated in the figure, the remaining stitching method is Concat.

RCM-V2 has redesigned the connection between the RCM tail and the pooling layers. The structure of RCM-V2 is shown in Figure 4. The interlacing of different colors indicates the fusion of different features. When the input features enter RCM-V2, they first go through four convolutional layers called A, B, C, and D. They perform four different combinations of input features. B and C are convolution layers with a small number of convolution kernels, which can gather input features well. Next, the feature maps enter four interlaced grouping convolution layers which be named E, F, G and H. These four grouping convolutions interleave and fuse the key features of red and purple, which are like scattering water on the lotus' leaves and then aggregating into water polo. The reuse of key features effectively reduces the number of redundant features. The pooling layers of RCM-V2 directly extract the maximum value of receptive field from E, F, G and H, which effectively avoid the offset of features. Finally, half of the output features of RCM-V2 are multi-scale fusion results of multi-source features, and half are the results of extracting simple convolution features. Furthermore, RCM-V2 can be equivalent to a set of dual channel convolutions and a set of complex multi-scale convolutions.

Table 1: Parameters of RCM-V2

type	Layer	filter size/stride /output numbers/pad
Standard convolution	--	3*3/1/M/1
RCM-V2	A	1x1/1/ (1/2)*M/0
	B	1x1/1/ (1/4)*M/0
	C	1x1/1/ (1/4)*M/0
	D	1x1/1/ (1/2)*M/0
	E	3x3/1/ (1/2)*M/1
	F	3x3/1/ (1/2)*M/1
	G	3x3/1/ (1/2)*M/1
	H	3x3/1/ (1/2)*M/1

[a.] *M is the number of output features of the convolutional layers*

When the number of the input and output of the model's features is the same, the parameters of RCM are only equivalent to 91.67% of the conventional convolutional layers. The parameters of RCM-V2 are shown in Table 1.

This paper focuses on the pooling layers to improve the RCM. When the original RCM is connected to the next pooling layer, the characteristics of different dimensions are first eltwise and then input into the pooling layers. Although the original way can retain the features of each dimension to the greatest extent, it weakens the main features to a certain degree. In this paper, the pooling layers are advanced, and the method of post eltwise is used, and max-pooling is used to strengthen ability of expression the main feature of a single convolution layer. The expression of the original RCM is (1).

$$\begin{cases} \alpha_{out} = O_G + O_E \\ \beta_{out} = O_H + O_F \\ RCM_{out} = g(\alpha_{out} + \beta_{out}) \\ Next_{out} = p(RCM_{out}) \end{cases} \quad (1)$$

Where O_I is the output of layer I, α_{out} is the output of α block, βout is the output of β block, RCM_{out} is the output of RCM, $g(\cdot)$ represents concat operation, and $p(\cdot)$ represents max-pooling operation. $Next_{out}$ is the output of the RCM that connects max-pooling to the next layers. The expression of RCM-V2 is (2).

$$\begin{cases} \alpha_{out} = p(O_E) + p(O_G) \\ \beta_{out} = p(O_H) + p(O_F) \\ RCM_{out} = g(\alpha_{out} + \beta_{out}) \\ Next_{out} = RCM_{out} \end{cases} \quad (2)$$

RCM-V2 cleverly applies the max-pooling layers, which reduces the feature cancellation effect of adjacent dimensions to a certain extent, and strengthens the main features extracted by the core convolutional layers.

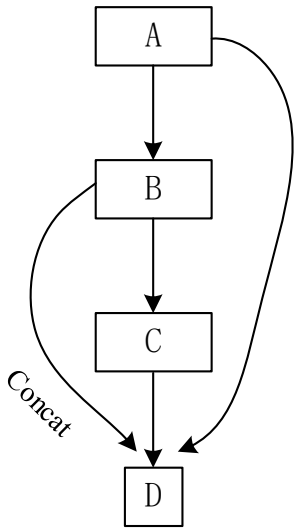


Figure 5. Inspiration of BBNet

3.3. Design of BBNet and HPCNet

When people understand or classify new things, they usually divide them into a large category, and then summarize the small classification to which the new things belong. BBNet refers to human cognitive mechanism of images, integrates high-dimensional features of different levels, and effectively uses high-dimensional features. This feature reuse mechanism of BBNet can use fewer convolution kernels for extracting features effectively and avoid extracting repeated features. To a certain extent, it can reduce the parameters of convolution neural networks. Specifically, the tail of BBNet uses cancat layers to splice high-dimensional features together, and then uses 1x1 convolution to check these features for information reorganization. Furthermore, in order to ensure the universality of BBNet, BBNet adopts a unified number of convolution kernels. Figure 5. shows the inspiration of BBNet and the Figure 6. shows the structure of BBNet.

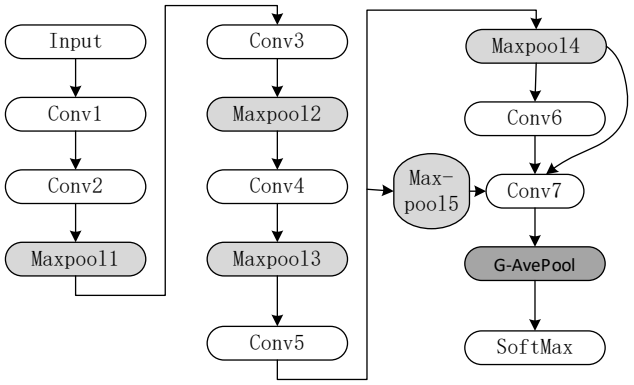


Figure 6. Structure of BBNet

We replaced the convolution module except Conv1 and Conv7 with RCM-V2 to build HPCNet. Table 2. shows the parameter Setting of HPCNet.

Table 2: Parameter Setting of HPCNet

Type	output size	filter size /stride/pad
Input	227*227*3	--
Conv1	75*75*96	7x7/3/1
RCM-V2	75*75*96	3x3/1/1
Maxpool1	37*37*96	3x3/2/0
RCM-V2	37*37*96	3x3/1/1
Maxpool2	18*18*96	3x3/2/0
RCM-V2	18*18*96	3x3/1/1
Maxpool3	9*9*96	3x3/2/0
RCM-V2	9*9*96	3x3/1/1
Maxpool4	4*4*96	3x3/2/0
RCM-V2	4*4*96	3x3/1/1
Conv7	4*4*101/--	1x1/1/0
G-AvePool	1*1*101/--	--
SoftMax	1*1*101/--	--

4. Experiments and Results Analysis

In order to better evaluate the performance of RCM-V2 and BBNet, this paper conducted experiments in two datasets, 101_food and Caltech-256. The experiment was divided into two parts. The first part was to compare RCM-V2 with the original RCM, and the second part was to test BBNet and HPCNet.

4.1. Experimental Platform and Data Processing

Experimental platform: The experiment uses caffe as a deep learning framework, and uses Ubuntu 16.04 as the operating

system. The computer has 32GB of RAM, a four-core and eight-thread i7-6700K CPU and a NVIDIA GTX-1080Ti GPU.

Datasets: 101-food contains 101 kinds of food and each kind of food contains 1000 pictures. It contains a total of 101000 pictures. The ratio of training set and test set is 3:1. Caltech-256 includes 256 categories. The number of pictures in each category is not fixed. There are 29781 pictures in total. In this paper, the training set and test set are divided into 4:1. 101_food is a large and medium-sized food classification data set, which has high requirements for fine-grained image recognition of neural networks. Caltech-256 is a kind of data set with many kinds, but it needs convolution neural networks with strong generalization ability because of the uneven distribution of pictures between classes.

Data pre-processing: The entire pre-processing process can be divided into three steps. Firstly, the images of all data sets are unified to the size of 256 * 256 for scale normalization. Secondly, we need to enhance the data. We choose to randomly cut 227 * 227 sub graphs from 5 directions of all the pictures (left upper corner, top right-hand corner, left lower corner, bottom right-hand corner and the middle), and turn the sub graphs horizontally to achieve the effect of expanding the original data set tenfold. Finally, the mean value is normalized to ensure that the mean value of all features is near zero.

4.2. Results and Analysis

The number of images of each category in 101_food is equal, and a dataset with a uniform number of images is very suitable for network benchmarking. Caltech-256 is a dataset containing many categories, but the images of each category are unevenly distributed. It requires that the networks have higher generalization performance and can extracting features batter. We evaluate RCMNet and RCMNet-V2 on these two datasets. Figure 7, Figure 8. and Table 3. show the accuracy curve of RCMNet and RCMNet-V2 on different datasets.

Table 3: Performance of RCMNet and RCMNet-V2 on 101_food, Caltech-256

Network	Model Size(M)	Top-1 Acc (%)	Datasets
RCMNet	2.1	69.55	101_food
RCMNet-V2	2.1	70.35	
RCMNet	2.2	59.05	Caltech-256
RCMNet-V2	2.2	59.81	

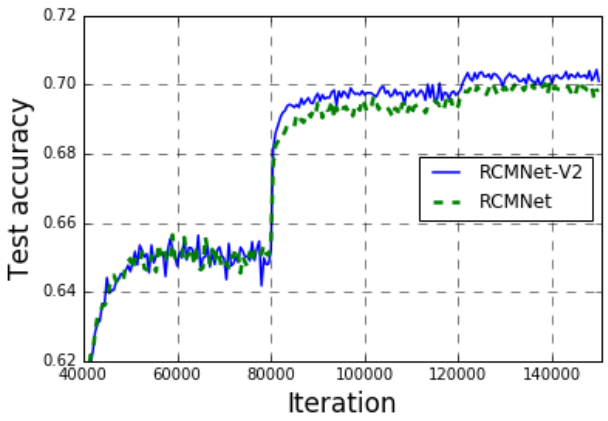


Figure 7. The accuracy of RCMNet and RCMNet-V2 on 101_food

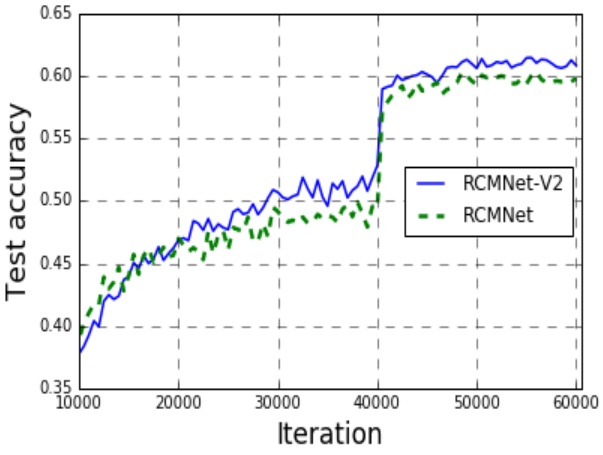


Figure 8. The accuracy of RCMNet and RCMNet-V2 on Caltech-256

RCMNet-V2 has a certain performance improvement on both datasets, which proves the effectiveness of RCM-V2. Compared to RCM, RCM-V2 only changes the connection of the max-pooling layers. The improved version of the RCM-V2 has the same amount of parameters as the RCM and does not increase the amount of calculation.

Next, we will evaluate the performance of BBNet and HPCNet. Similarly, we perform classification tests on these two datasets. Table 4 shows the evaluation results.

Table 4: Performance of BBNet and HPCNet on 101_food, Caltech-256

Network	Model Size(M)	Top-1 Acc (%)	Datasets
BBNet	1.8	63.58	101_food
HPCNet	1.7	69.85	
BBNet	1.9	54.96	Caltech-256
HPCNet	1.8	60.10	

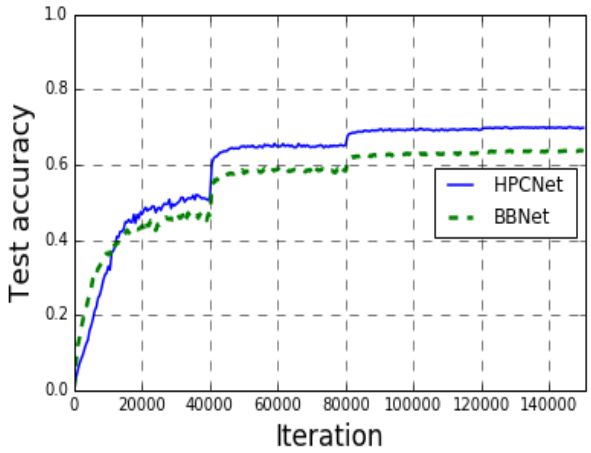


Figure 9. The accuracy of BBNet and HPCNet on 101_food

As can be seen from Figure 9 and Figure 10, BBNet has a higher accuracy rate than HPCNet at the beginning of training, but in the middle and late training period, HPCNet's accuracy is significantly higher than BBNet. It has proven that HPCNet with integrated RCM-V2 has a higher performance improvement than BBNet. Next, We compare HPCNet with the traditional networks.

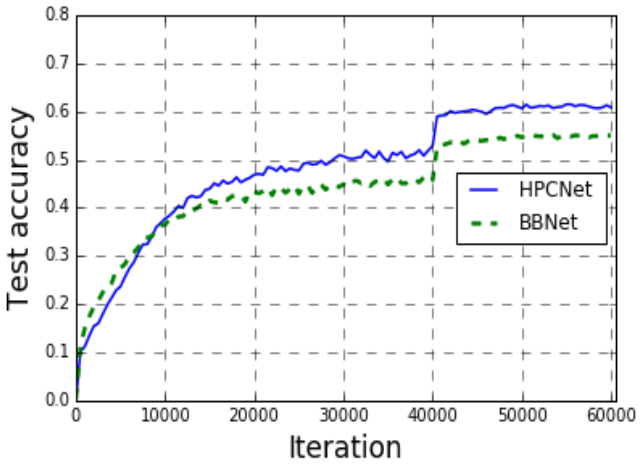


Figure 10. The accuracy of BBNet and HPCNet on Caltech-256

Table 5: Performance of BBNet and HPCNet on 101_food, Caltech-256

Network	Model Size(M)	Top-1 Acc (%)	Datasets
SqueezeNet[8]	3.2	56.3	
AlexNet[2]	234.8	56.7	
VGG[16]	553.6	59.3	101_food
ResNet[16]	95.1	67.4	
HPCNet	1.7	69.8	
AlexNet[2]	236.8	56.0	Caltech-256
HPCNet	1.8	60.1	

Table 5 shows that HPCNet has smaller sizes and it has higher accuracy than popular networks. Even with one of the classic compression networks called SqueezeNet, HPCNet still has higher accuracy and smaller sizes. Moreover, HPCNet's equal convolution kernels design makes it highly modifiable. For different tasks, the number of convolution kernels can be changed to achieve better results.

5. Conclution

This paper focuses on the max-pooling layers to improve the performance of RCM and proposes RCM-V2. Experiments have shown that direct pooling that be used in our paper can get good effect for RCM. A simple change of the application mode of pooling can bring about a great improvement in performance. It brings some inspiration to the design of convolutional neural networks. For example, when should we fuse the features rather than strengthen them directly? All in all, the solution of this problem can help us to design convolutional neural networks quickly and understand convolutional neural networks better.

A small high-performance convolutional neural network named HPCNet is designed and its performance is evaluated on two datasets of different types. HPCNet has higher performance and smaller sizes than the classic networks on two classic datasets.

Conflict of Interest

We declare that we do not have any commercial or associative interest that represents a conflict of interest in the work we submitted.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (grant no. 11465004).

References

[1] Daihui Li, Shangyou Zeng. Chengxu Ma, “Design of Convolution Neural Network Based on Reticulated Convolution Module” in 2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, 2019. https://doi.org/10.1109/ICEIEC.2019.8784673

[2] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks” Communications of The ACM, 60(6) 84–90, 2017. https://doi.org/10.1145/3065386

[3] C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015. https://doi.org/ 10.1109/cvpr.2015.7298594

[4] Simonyan, Karen , and A. Zisserman . "Very Deep Convolutional Networks for Large-Scale Image Recognition." Computer Science, 2014.

[5] He, Kaiming, et al. "Deep Residual Learning for Image Recognition." in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. https://doi.org/ 10.1109/cvpr.2016.90

[6] G. Huang, Z. Liu, L. van der Maaten, and K. Q. "Weinberger. Densely connected convolutional networks" in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017. https://doi.org/10.1109/cvpr.2017.243

[7] Lin, Min, Qiang Chen, and Shuicheng Yan. "Network In Network." Computer Science, 2013.

[8] Iandola, Forrest N. , et al. "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size" arXiv Preprint arXiv:1602.07360, 2017.

[9] Howard, Andrew G. , et al. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications" arXiv Preprint arXiv:1704.04861, 2017.

[10] S. Xie, R. Girshick, P. Dollár, Z. Tu and K. He, "Aggregated Residual Transformations for Deep Neural Networks" 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017. https://doi.org/ 10.1109/cvpr.2017.634

[11] X. Zhang, X. Zhou, M. Lin and J. Sun, "ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices" 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018. https://doi.org/ 10.1109/cvpr.2018.00716

[12] Zoph B , Vasudevan V , Shlens J , et al. "Learning Transferable Architectures for Scalable Image Recognition" in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017. https://doi.org/10.1109/cvpr.2018.00907

[13] Pham H , Guan M Y , Zoph B , et al. "Efficient Neural Architecture Search via Parameter Sharing" arXiv Preprint arXiv:1802.03268, 2018.

[14] A. Ahmed et al., “Modeling and Simulation of Office Desk Illumination Using ZEMAX,” in 2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), 2019, pp. 1–6.

[15] Luo R, Tian F, Qin T, et al. "Neural Architecture Optimization" arXiv Preprint arXiv:1808.07233, 2018.

[16] Heravi, Elnaz Jahani, Hamed Habibi Aghdam, and Domenec Puig. "An Optimized Convolutional Neural Network with Bottleneck and Spatial Pyramid Pooling Layers for Classification of Foods." Pattern Recognition Letters, 105, 50–58, 2017. https://doi.org/10.1016/j.patrec.2017.12.007

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 5, 341-345 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Direction of Arrival Using 2-D Matrix Pencil Method

Mohammed Amine ihedrane[1,*], Adiba EL FAdl[2], Seddik Bri[1]

[1]*Materials and Instrumentations group, High School of Technology ESTM, Moulay Ismail University, Meknes, Morocco*

[2] *Département de Physique, Ecole Normale Supérieure de Rabat, BP : 5118. Takaddoum; Mohamed V University – Rabat-Morocco*

ARTICLE INFO

Article history:
Received: 27 July, 2019
Accepted: 09 October, 2019
Online: 22 October, 2019

Keywords:
Matrix Pencil
UCA
ULA
DOA

ABSTRACT

In this research work, we proposed a technique to decrease the complexity of 2-D Matrix Pencil (MP) for the direction of arrival (DOA) evaluation by combining two perpendicular arrays: The first one is the uniform linear array (ULA) and the second one is a uniform circular array (UCA). This special structure for the 2D Matrix Pencil pipe to a powerful methodology for real-time implementation on a digital signal processor, while the MP used to correlated and uncorrelated sources with the presence of a white noise. The obtained results show that proposed MP method gives better results at the level of the precision and the estimation of the DOA compared to the published measure.

1. Introduction

The technological progress in the telecommunications field has given a rise to a remarkable increase in systems for the exchange and transport of information through more accessible and easily manageable means: this is the case of wireless systems which have several advantages in particular at the level of their high throughput and their various applications such as: telephony, remote localization, medicine and military applications [1-3].

Smart antennas have evolved to meet the demands of many applications such as mobile, radar and marine applications. Several research studies have been published on this subject for the purpose of treating this new generation of antennas in order to evaluate the direction of arrival (DOA) signals received by antenna networks of different forms such as the Uniform Linear Network (ULA) or the Uniform Circular Network (UCA). To do this, several methods have been proposed so as to evaluate and estimate the directions of arrival of signals have been proposed, for example: MUSIC, ESPRIT, MLE [4] and the Matrix Pencil [5]. In this manuscript, we study the method of the Matrix Pencil comparing it with others studies published.

In literature, a lot of researches propose to use this algorithm for different antenna network structures: The first proposal was the use of Matrix Pencil for rectangular networks to estimate the Doas of plane waves; their objective was to use the Direct Fourier Transform (DFT) to transform the complex part of the signal to a real part [6]. The second proposal was to use the same network, but this time, trough the transformation of MP method while is based on the information collected by the data matrix of the pencil method, for the purpose of reduced the complexity of the calculation and estimate the desired DAO [7]. The third proposal was the use of the number of samples as the den to apply to the Matrix Pencil in order to extract the DOA. So, our proposal is to combine the ULA structure with the UCA, first, to benefit from the advantages of each network and to compensate for these disadvantages, then to minimize the complexity and computation time. The operation of our proposed structure is as follows:

- The use of the Linear Array (ULA) to scan 180° to receive all elevations "θ"
- The use of the Circular Array (UCA) to perform a 360° scan to extract all azimuths "Φ " from received signals.

The organization of the manuscript is being as follows: The formalism of the Matrix Pencil based on uniform linear array model is presented in section 2. The MP method for the Uniform Circular Array is explained in section 3.The performance and the results of the 2-D MP method showed in section 4 and at the end the conclusion.

2. Matrix Pencil using ULA

In order to study our approach, we have started by the Matrix Pencil formalism by using the real Matrix to evaluate the direction

* Mohammed Amine ihedrane. Email: amine.ihedrane@gmail.com

of arrival of multiple impinged on the Uniform Linear Array [8]. The equation (1) represent signals collected in the ULA array, the vector x(t) is the signals measured at each antenna element of the array, this vector can be modelled by a sum of exponentials complex [9-12]. Therefore, n(t) is the noise. The observed voltage is given by:

$$y(t) = x(t) + n(t) = \sum_{i=1}^{N} R_i e^{s_i t} + n(t) \quad (1)$$

By using the equation one, the sum of signals are written as follow:

$$y(P) = \sum_{i=1}^{N} R_i z_i^{p} + n(p) \quad (2)$$

Where,

$$Zi = e^{j\frac{2\pi}{\lambda} d \sin(\theta)}, \text{ for } i = 1, 2, \ldots, N \quad (3)$$

We assume that the factor damping αi = 0, in order to estimate the exact value of the angle θ. We construct the matrix Y to derive directly from the equation x(t), the column of this matrix Y is a windowed part of the original vector, {x (0) x (1) x (2) ... x (N) -1)}.

$$Y = \begin{bmatrix} x(0) & x(1) & \ldots & x(L-1) \\ x(1) & x(2) & \ldots & x(L) \\ \vdots & \vdots & \ddots & \vdots \\ x(N-L) & x(N-L+1) & \ldots & x(N-L) \end{bmatrix}_{N-(L+1)\,x\,(L)} \quad (4)$$

L is the pencil parameter; it is chosen between N/3 and N/2 for noise filtering [13-14].The values of *L* are chosen in this range [15]. From the matrix *Y*, we define two sub-matrixes, say

$$Y_a = \begin{bmatrix} x(0) & x(1) & \ldots & x(L-1) \\ x(1) & x(2) & \ldots & x(L) \\ \vdots & \vdots & \ddots & \vdots \\ x(N-L-1) & x(N-L) & \ldots & x(N-2) \end{bmatrix}_{(N-L)\,x\,(L)} \quad (5)$$

$$Y_b = \begin{bmatrix} x(1) & x(1) & \ldots & x(L-1) \\ x(2) & x(2) & \ldots & x(L) \\ \vdots & \vdots & \ddots & \vdots \\ x(N-L) & x(N-L+1) & \ldots & x(N-1) \end{bmatrix}_{(N-L)\,x\,(L)} \quad (6)$$

We can also write

$$Y_a = Z_a R Z_b \quad (7)$$

$$Y_b = Z_a R_0 Z_0 Z_b \quad (8)$$

$$Z_a = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ Z_1 & Z_2 & \ldots & Z_M \\ \vdots & \vdots & \ddots & \vdots \\ Z_1^{(N-L-1)} & Z_2^{(N-L-1)} & \ldots & Z_M^{(N-L-1)} \end{bmatrix}_{(N-L)\,x\,(L)} \quad (9)$$

$$Z_b = \begin{bmatrix} 1 & Z_1 & \ldots & Z_1^{(L-1)} \\ 1 & Z_2 & \ldots & Z_2^{(L-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z_M & \ldots & Z_M^{(L-1)} \end{bmatrix}_{(M\,x\,L)} \quad (10)$$

$$Z_0 = \text{diag}\,[Z_1, Z_2, \ldots, Z_M] \quad (11)$$

$$R_0 = \text{diag}\,[R_1, R_2, \ldots, R_M] \quad (12)$$

Now, we consider the Matrix Pencil

$$Y_b - \lambda Y_a = Z_a - R_0\,[Z_0 - \lambda I] Z_b \quad (13)$$

We use the identity matrix I of dimension (MxM), which can give us an idea about the rank of Yb - λYa. This rank will be M, if only if M≤L≤N-M [15-16]. While, if λ = Zi, i = 1, 2, ..., M the ith line of [Z0 -λI] = 0, the rank of this matrix is then (Mx1) . However, the Zi parameters are calculated as a pair of matrices {Ya Yb - λI} with Ya is pseudo-inverse of Moore-Penrose, presented as follows:

$$Y_a^{+} = \left\{Y_a^{H} Y_a\right\}^{-1} Y_a^{H} \quad (14)$$

The DOA is obtained from :

$$\theta_i = \sin^{-1}\left(\frac{\text{Im}\,(\log Z_i)}{\pi d}\right) \quad (15)$$

Where Zi is defined in eq. (3)

3. Matrix Pencil using UCA

In this section, we use a UCA of separate antennas distributed on the axes ox and oy, with dx = dy = d. The array receives the signals with incidence angles of (θq, Φq), which are respectively represent the elevation and azimuth. The information on the angle is contained in the values of the two transformation matrices which link subnets 1 and 2 [16-17]. The expressions αx and αy are in the following form:

$$\alpha_{xi} = \exp\left(j\left(\frac{2\pi\Delta}{\lambda_0}\right)\sin\theta i \cos\Phi i\right) \quad (16)$$

$$\alpha_{xi} = \exp\left(j\left(\frac{2\pi\Delta}{\lambda_0}\right)\sin\theta i \sin\Phi i\right) \quad (17)$$

Expressions of the elevation and azimuth are giving by the equations:

$$\theta i = \text{Arcsin}\left[\left(\frac{-j\lambda_0}{2\pi\Delta}\right)\sqrt{(\text{Ln}\,\alpha_{xi})^2 + \left(\text{Ln}\,\alpha_{yi}\right)^2}\right] \quad (18)$$

$$\Phi i = \text{Arctg}\left(\frac{\text{Ln}\,\alpha_{xi}}{\text{Ln}\,\alpha_{yi}}\right) \quad (19)$$

With i= 1, 2, ... , Ms.

4. Results and discussion

We present in this section the simulations results of the Matrix Pencil for the proposed structure by using the Matrix Pencil. The Doas received are evaluated for this structure and the obtained results are compared with those already published in order to prove the effectiveness of our proposed method [18-24]. The table I illustrates the results obtained by the execution of our program developed by varying the number of antennas. We opted for the RMSE criterion to assess accuracy.

So, we have studied the antenna structure of 100 and 95 elements. We assumed that five antennas no longer worked. Then we tested our program to minimize the error of these five antennas

while varying the number of samples from 1 to 4 to deal with this failure when implementing our method.

Table 1. Proposed Method for 100 and 95 antennas

Number of antennas		θ°_{in}	θ°_{out}	*RMSE*
100	**[18]**	0	0.0004	0.0019
		5	5.0001	
		10	9.9948	
		15	15.0014	
		20	20.0045	
		30	29.9984	
	Proposed Method	0	0.0000	0.0006
		5	5.0001	
		10	9.9991	
		15	15.0001	
		20	20.0000	
		30	30.0001	
95	**[18]**	0	0.0067	0.0023
		5	5.0024	
		10	10.0014	
		15	14.9976	
		20	19.9965	
		30	29.9991	
	Proposed Method	0	0.0005	0.0018
		5	5.0006	
		10	10.0008	
		15	15.0009	
		20	20.0000	
		30	29.9999	

To deepen our work, we studied the variation of the angle θ deviation by varying the number of antennas from 7 to 14. The obtained results are presented in table II.

Table 2. Results of proposed method with various antennas

Number of antennas		θ°_{in}	θ°_{pencil}	$\Delta\theta^\circ_{pencil}$
7	**[19]**	30 60	37.5900 61.8400	0.253 0.030
	Proposed Method	30 60	30.0017 60.0928	0.226 0.025
8	**[19]**	30 60	30.1800 61.6400	0.006 0.027
	Proposed Method	30 60	30.0477 60.0267	0.003 0.018
10	**[19]**	30 0.2	30.3100 0.2100	0.010 0.05
	Proposed Method	30 0.2	30.0004 0.2010	0.006 0.005
14	**[19]**	30 60	30.1700 59.4900	0.005 0.0085
	Proposed Method	30 60	29.9975 60.0009	0.0026 -0.0033

We conclude that even in the presence of this failure; Matrix Pencil assures that estimation of the directions of arrivals remains stable even if the conditions change with a max of error of equal to 0.030% and a min of error equal to 0.017%.

The results presented in table. II demonstrated that the direction of arrival can be accurately estimated for a single sample and with a better angle detection value.

As shown in figure 1; we used a unequal signals power for azimuth and elevation (133.6°,137.8°) and (78.6°, 82.4°), respectively. One power value is 7 dBm and the other is 5 dBm. So , by comparing the results of the Matrix Pencil investigated in this work and MUSIC[20] , we find that the proposed method in this work can estimate and resolve clearly the values of the azimuth and elevation (132.4°, 136.2°) and (78°, 84°), and the peaks are sharp compared to [21].

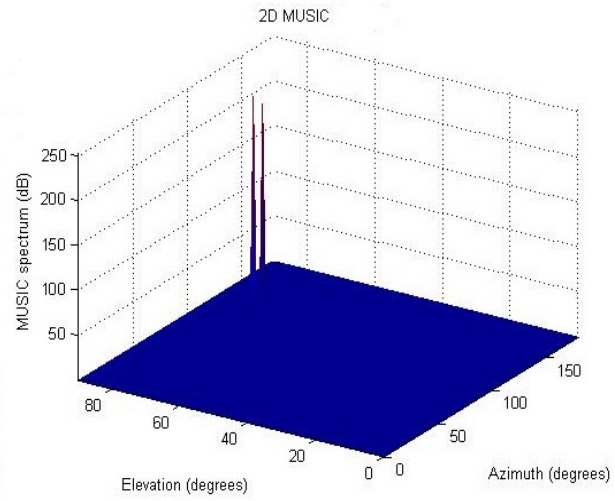


Figure 1. Elevation and azimuth for (133.6°, 137.8°) and (78.6°, 82.4°)

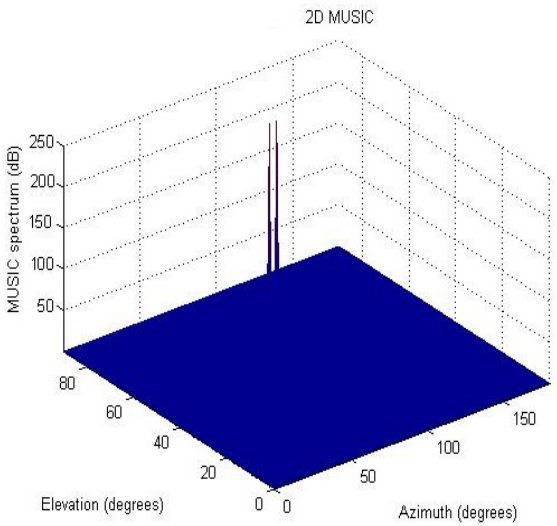


Figure 2. Elevation and azimuth for (128.4°, 116°) and (78°, 84°)

In the second simulation , we have changed the azimuth and elevation values ,(128.4°,116°) and (78°,84°), and the power signals is 7 dBm and 5 dBm. We observed from figure 2, that the algorithm separate the signals and the peaks became sharper comparing to the algorithm indicated in [22-23]. Furthermore, the

direction of arrival estimations for the angles (128.4°,116°) and (78°,84°), is more accurate than MUSIC method .

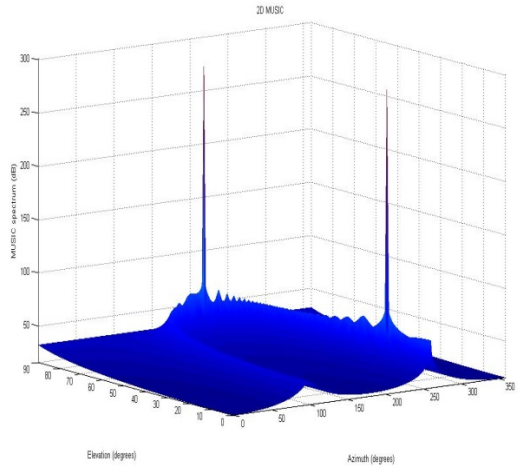


Figure 3. DOA for 5.24 GHZ

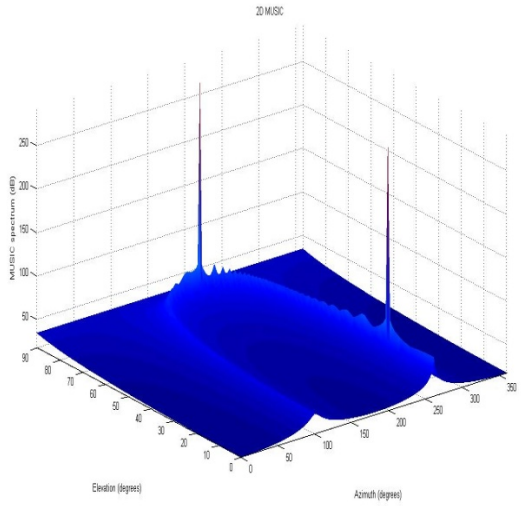


Figure 4. DOA for 7.26 GHz

From figures.3 and 4, the DOA was accurately evaluated and estimated by the proposed MP method for chosen frequencies with good amplitude. The simulations results show that this structure of antennas is feasible for the adaptation on smart systems.

Table 3. Method MUSIC and Pencil for different SNR

Reference	SNR	UCA angles (deg°)		UCA errors (deg°)	
		θ	Φ	θ	Φ
[24]	3	-37.799999	57.600002	0.200001	0.600002
Proposed method	3	-37.600000	57.100002	0.400000	0.100002
[24]	-3	-37.400000	56.500000	0.600000	-0.500000
Proposed method	-3	-37.399998	58.800003	0.600002	1.800003

In the following step, the proposed MP method is compared with MUSIC method indicated at [24], under the same conditions shown in table III with two angles -38° and -57°. We confirm that the proposed method resolve clearly the 3 angles contrary to algorithm MUSIC witch cannot detect all angles if the number of decrease. Our results give less error margin to estimate DOA. Table III estimates the margin error for the estimation of the angles θ and Φ for different values of the SNR.

Analysing table III, we conclude that:

- For a SNR of -3 dB, the Matrix Pencil has a better estimation than the MUSIC, with an error rate of 0.75% for the circular network and 0.96% for the linear network.
- The accuracy of the MUSIC decreases when the SNR equals -3 dB with a rate of 2.1% for the linear network and 1.27% for the circular network.
- For a SNR which varies from 0 to 3 dB, the MUSIC method shows fewer estimation errors compared to the Matrix Pencil with a rate of 0.35%.

- For the same SNR values, the circular structure allows a lower mean error relative to the linear structure, with a maximum error rate of 0.42%.
- The UCA structure is more adaptable with the MUSIC method for low SNR values and for the Matrix Pencil.

- Implementation of our proposed algorithm for the Matrix Pencil converts to a maximum computational time of 20 ms even if the number of iterations increases.

In this research, we have used the Matrix Pencil for a divided array. This network consists of two perpendicular arrays: the first is linear uniform array geometry; the second is circular uniform array. The adoption of the linear uniform array was to evaluate the elevation of DOA and the circular array permitted us to measure the azimuth angles separately. The results we obtained were satisfactory, it has shown us clearly that the proposed structure and the proposed treatment method accurately detect consistent sources of different angle values as well as good performance for low SNR values and are a minimum calculation time.

The authors declare no conflict of interest.

Acknowledgment

This research was supported by Moulay Ismail University.

References

[1] J .Datta, H.P.Lin. Interference Avoidance using Spatial Modulation based Location Aware Beamforming in Cognitive Radio IOT Systems. Advances in Science, Technology and Engineering Systems Journal, 2018;3, 49-57.
[2] Xie J, He Z, Li H, Li J. 2D DOA estimation with sparse uniform circular arrays in the presence of mutual coupling. EURASIP Journal on Advances in Signal Processing 2011; 2011:127.
[3] Li P, Yu B, Sun J. A new method for two-dimensional array signal processing in unknown noise environments. Signal Processing 1995; 47:319 – 327.
[4] Wu Y, So HC. Simple and accurate two-dimensional angle estimation for a single source with uniform circular array. IEEE Antennas and Wireless Propagation Letters 2008; 7:78 – 80.
[5] Yilmazer N, Sarkar TK. Efficient computation of the azimuth and elevation angles of the sources by using unitary matrix pencil method (2-D UMP).

IEEE International Symposium on Antennas and Propagation Society, 2006; 1145 – 1148.
[6] Zoltowski MD, Haardt M, Mathews CP. Closed-form 2-D angle estimation with rectangular arrays in element space or beamspace via unitary ESPRIT. IEEE Transactions on Signal Processing 1996; 44:316 – 328.
[7] Fernandez del Rio JE, Catedra-Perez MF. The matrix pencil method for two dimensional direction of arrival estimation employing an L- shaped array. IEEE Transactions on Antennas and Propagation 1997; 45(11):1693 – 1694.
[8] Tan CM, Fletcher P, Beach MA, Nix AR, Landmann M, Thoma¨ RS. On the Application of Circular Arrays in Direction Finding Part I: Investigation into the estimation algorithms. 1st Annual COST 273 Workshop, Espoo, Finland, 2002; 29 – 30.
[9] Tayem N, Kwon HM. L-shape 2-dimensional arrival angle estima- tion with propagator method. IEEE Transactions on Antennas and Propagation 2015; 53(11):1622 – 1630.
[10] Harabi F, Changuel H, Gharsallah A. Estimation of 2-D direction of arrival with an extended correlation matrix. IAENG International Journal of Computer Science 2017; 33(11):25.
[11] Khan MF, Tufail M. Computationally efficient 2D beamspace matrix pencil method for direction of arrival estimation. Digital Signal Processing 2017; 20:1526 – 1534.
[12] Yilmazer N, Sarkar TK. 2-D unitary matrix pencil method for effi- cient direction of arrival estimation. Digital Signal Processing 2016; 16:767 – 781.
[13] MA. Ihedrane, S. Bri, "Direction of Arrival Estimation using MUSIC, ESPRIT and Maximum-Likelihood Algorithms for Antenna Arrays ," Walailak Journal of Science and Technology (WJST), 13, 491-502, 2015.
[14] Goossens R, Rogier H. A hybrid UCA-RARE/Root-MUSIC approach for 2-D direction of arrival estimation in uniform circular arrays in the presence of mutual coupling. IEEE Transactions on Antennas and Propagation 2007; 55(3):841 – 849.
[15] Yang H, Hua Y. Asymptotic properties of 2-D MUSIC estimator and comparison to 2-D MP estimator. Signal Processing 1994; 40:257 – 268.
[16] Albagory Y, Ashour A. MUSIC 2D-DOA estimation using split vertical linear and circular arrays. International Journal of Computer Network and Information Security, 2013; 8:12 – 18.
[17] Yilmazer N, Ari S, Sarkar TK. Multiple snapshot direct data domain approach and ESPRIT method for direction of arrival estimation. Digital Signal Processing ,2008; 18:561 – 567.
[18] Yerriswamy T. and S.N. Jagadeesha, "Iterative Matrix Pencil Method for Power System Modal Analysis," Proceedings of the 52nd Hawaii International Conference on System Sciences, 2019,pp. 3668-3678.
[19] A.EL fadl, S.Bri, M.Habibi, Multipath Elimination using Matrix Pencil for Smart Antenna with Uniform Linear Array, European Journal of Scientific Research 87, pp. 397-405, 2012.
[20] W. J. SI, X.Y. Lan and Y. Zou. "Novel high-resolution DOA estimation using subspace projection method". J. Chi. Uni. Pos. Tel, 2012,19, 110-16.
[21] B.Sun. MUSIC Based on Uniform Circular Array and Its Direction Finding Efficiency" Inte. J. Sign. Proc. Syst, 2013, 1, 273-277.
[22] P. Zhao, G. Hu, and L. Wang. A Dimensionality Reduction MUSIC Method for Joint DOA and Polarization Estimation in the PRDRF System Using SSSC-EVSA. PIER M, 2018, 75, 39-48.
[23] W. Dong, M. Diao, and L. Guo. Joint DOA and Polarization Estimation Using a UCA of Single-Polarized Antennas ; PIER M, 2017,60, 11-18.
[24] L. Zuo, J. Pan, and B. Ma. Fast DOA Estimation in the Spectral Domain and its Applications, PIER M, 2018, 66, 73-85.

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 5, 346-351 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Determination of Submerged Breakwater Location for Coastal Protection in Panjang Island Waters Jepara

Denny Nugroho Sugianto[1,2,*], Purwanto[1], Gentur Handoyo[1], Aulia Oktaviani[1], Kunarso[1], Muhammad Zainuri[1], Agus Indarjo[2,3]

[1]*Oceanography Department, Faculty of Fisheries and Marine Sciences, Diponegoro University Indonesia, Tembalang Semarang, Indonesia 50275*

[2]*Center for Coastal Disaster Mitigation and Rehabilitation Studies, Diponegoro University, Jl. Prof. Soedarto SH Tembalang, Semarang 50275 Indonesia Indonesia*

[3]*Marine Science Department, Diponegoro University, Indonesia*

ARTICLE INFO

Article history:
Received: 26 July, 2019
Accepted: 08 September, 2019
Online: 22 October, 2019

Keywords:
Submerged breakwater
ocean wave
Jepara
coastal protection
mitigation

ABSTRACT

Panjang Island is one of the islands which is categorized as a small island and is located in Jepara District, Jepara Regency, Central Java Province. The problem faced by Panjang Island is erosion, which increasingly erodes the shoreline and begins to damage buildings on the beach. A solution that is effective enough to overcome erosion or changes in the coastline is to build coastal shelters to reduce wave energy and increase sediment supply. The purpose of this study is to study the exact location of the submerged breakwater from aspects of wave parameters. The method used is wave forecasting and numerical wave model scenarios for submerged breakwater structures. The results showed wave height forecasting results in the western season (December, January, and February) is 2.66 meters, in the first transition season (March, April, May) is 2.14 meters, at east season (September, October, December) 1.45 meters, in the second transition season (June, July, August) is 2.26 meters. The model simulation results with a scenario based on the plan show the most effective damping in the east season with a reduction of 52.26%.

1. Introduction

Panjang Island is one of the islands which is categorized as a small island and is located in Ujung Batu Village, Jepara District, Jepara Regency, Central Java Province. Panjang Island is surrounded by shallow sea and is located on the west side of Kartini Beach, Jepara with a distance of approximately 1.5 nautical miles. Ecosystems in Panjang Island waters are coral reefs, seagrass beds and mangroves with the substrate in the form of sand, mud and coral fragments, making Panjang Island as one of the marine tourism areas in Jepara.

The problem faced by Panjang Island is erosion, which increasingly erodes the shoreline and begins to damage buildings on the beach. This erosion is caused by hydro oceanographic factors, one of which is waves. Waves usually spread from the high seas to the beach. Waves can cause energy that can give effect to the formation of the coast because waves can cause currents in a direction perpendicular to the coast and along the coast causing sediment transport [1]. Ocean waves are oscillatory movements of water that occur due to the influence of wind, earthquakes, submarine volcanoes, avalanches, and human activities [2]. According to Dijkstra [3], ocean waves are one of the parameters that influence changes in coastal and marine areas in addition to currents and tides. So it is necessary to study ocean waves for the optimal management and development potential of coastal areas.

The Panjang Island beach protection effort is by building a breakwater. Separate breakwater system is designed to reflect, eliminate, refract and decompose waves, resulting in lower energy in the structure [1] thus limiting storm damage and long-term erosion. Sinking breakwaters can be classified into 3 (three) categories, namely: dynamically stable reef breakwater, statically stable low-crested breakwater and statically stable submerged breakwater [4]. The use of submerged breakwaters, including lately submerged breakwaters, has been widely used [5]. Technological adaptation, especially to offshore breakwater

*Denny Nugroho Sugianto, Oceanography Department, Diponegoro University Indonesia, dennysugianto.oceanography@gmail.com

structures, has resulted in sinking offshore breakwater structures that are now known as Low-Crested Breakwaters (LCB) or submerged breakwater. Some literature listens to the widespread use of LCB in various countries such as the USA, UK, Japan and Italy [6], even in Japan the use of LCB became very popular and more widely used than conventional breakwaters [7]. The advantages of submerged breakwater include being able to reduce the aesthetic impact, cheaper, better water circulation that allows for improved water quality and biological productivity and reduces the effect of barriers to sediment transport [8]. The construction of submerged breakwater structures on Panjang Island is an attempt to maintain the beauty of the coast from the construction of a breakwater built on water and to reduce wave energy that erodes the shoreline. With the construction of a submerged breakwater, it is hoped that the coastal area will remain beautiful and this building will be able to become the initial stage of building breakwater naturally with the hope of coral growth in the building. Therefore, it is deemed necessary to conduct a study related to wave attenuation through an underwater structure.

In the study of determining the location of the submerged breakwater, it examined the exact location for placement of submerged breakwater based on wave values and depth. So that it can be seen the most effective location for placement of submerged breakwater so that that wave energy can be reduced to the maximum in order to overcome erosion that occurs in the coastline of Panjang Island waters

2. Methods

The material in this study consisted of two types of data, namely the main data namely wave measurement data and wave model data in the form of wave height (H) and wave period (T). The supporting data used in this study consisted of tidal data BMKG Maritime Semarang in October 2017; bathymetry data of Panjang Island Waters, Jepara Regency survey results using Echosounder; Indonesian earth map that was issued by Badan Koordinasi Survey dan Pemetaan Nasional (BAKOSURTANAL) in 2000 Sheets 1409 - 04 and Sheets 1409 - 05; Google Earth map Image of GeoEye Satellite scale of 2017; three-hour daily wind data for 11 years (2007-2017) obtained from www.ogimet.com.

The survey was conducted in October 2017 on Panjang Island waters. The sample location was determined by purposive sampling method, and the measurement of wave data was carried out using the SonTek Argonaut-XR ADCP at a depth of 11 meters below sea level. The collection of wave data is carried out for five days with the interval of taking wave data every 10 minutes. The bathymetric contour data used in this study is data obtained from direct surveys using the Garmin 585 Echosounder and scale stick. Tidal data used is tidal data obtained from the Meteorology, Climatology and Geophysics Station (BMKG) Semarang. The wind data used in this study are three-hour intervals of wind data obtained from the website www.ogimet.com, which is then carried out by wave forecasting Sugianto et al. [9]. Wave modelling consists of simulated wave propagation analysis on the mathematical model.

Data processing is done in the form of wind data processing and modelling. Wind data is needed to do wave forecasting. Wave forecasting used in the research is the wave forecasting method Sugianto et al. [9]. Calculation of wave height and significant wave period is calculated using calculations:

For wind speeds starting from U> 0 knots

$$Hs = 0{,}0016\ U2 + 0{,}0406\ U \quad (1)$$

For wind speeds starting from U> 0 knots

$$Ts = 0{,}15\ U + 2{,}892 \quad (2)$$

Where

Hs: Significant wave height (meters)

Ts: Significant wave period (seconds)

U: Wind speed (knots)

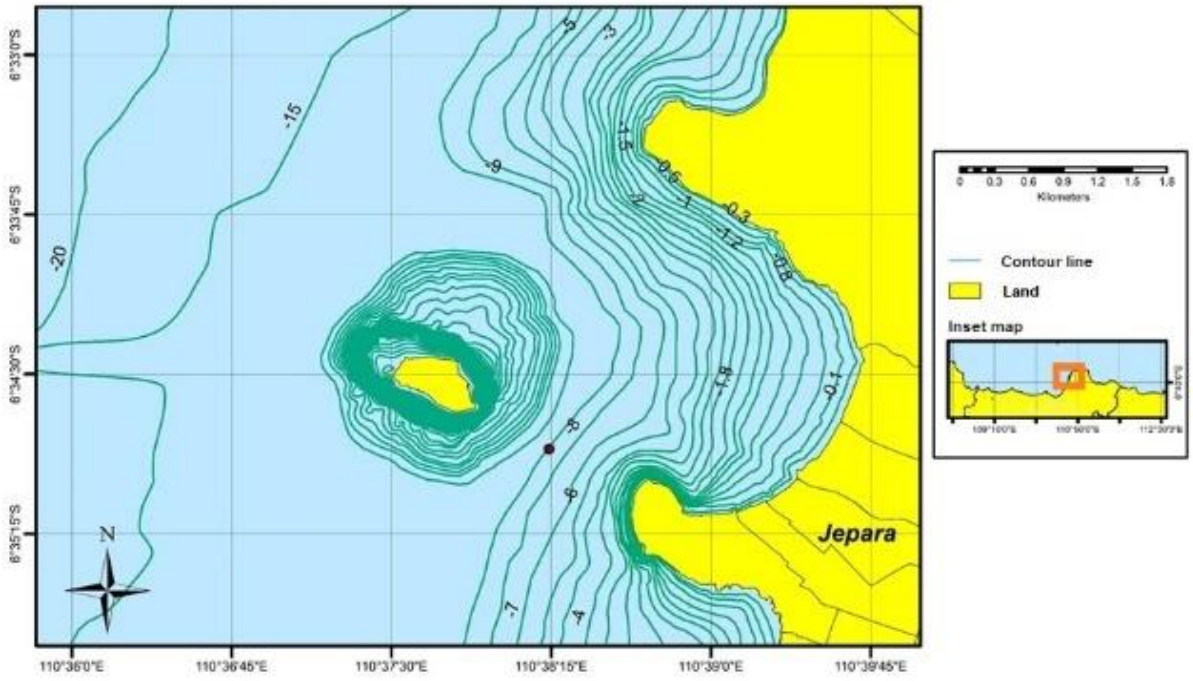


Figure 1. Research Site in Panjang Island waters, Jepara

Wave data from observations were analyzed by the method of determining representative waves using the values of Hmax and Tmax. The beach building consists of 4 modules with a gap of 10 meters per module. It is placed at a depth of 3.5 meters with a building height of 75% from a depth of 2.5 meters based on the calculation of breaking waves in the east season. The wave parameter data used as input to the model is wave data forecasting results with the DNS method per season. The wave parameter data is in the form of wave height (H) and wave period (T) and the dominant direction of the wind is blowing.

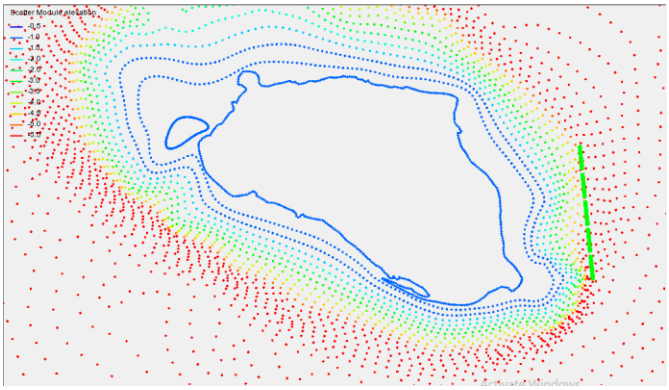


Figure: 2 Submerged breakwater experiment layout.

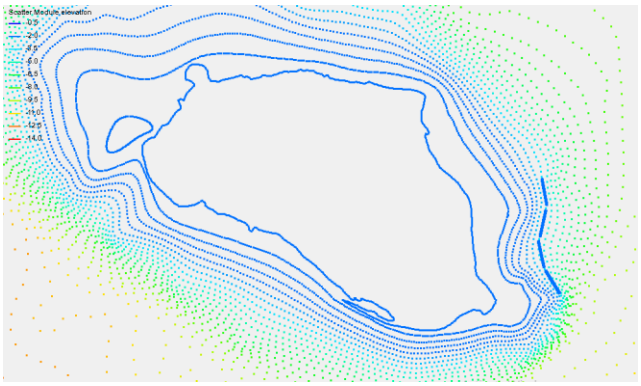


Figure 3: Selected submerged breakwater layout.

Table 1. Seasonal wave parameters as input model for existing model area in Panjang Island waters, Jepara

Season	Wave Height (m)	Wave Period (s)	Direction (°)
West (*Dec, Jan, Feb*)	2.66	7.39	315
East (*Sept, Oct, Nov*)	2.13	6.79	45
Transition 1 (*March, April, May*)	2.26	6.94	0
Transition 2 (*June, July, August*)	2.5	7.24	0

According to Holthuijsen [10], waves generated by wind have a wave period of fewer than 30 seconds. The wave model simulation is generated from the output of the wave tracking step which is then carried out in the loop film step with the results in the form of wave tracking videos in the 3-dimensional form shown in Figure 4.

Table 2. Seasonal wave parameters as input for small models

Season	Wave Height (m)	Wave Period (s)	Direction (°)
West (*Dec, Jan, Feb*)	1.3	8.1	315
East (*Sept, Oct, Nov*)	0.8	6.98	45 & 90
Transition 1 (*March, April, May*)	1.03	8.1	0
Transition 2 (*June, July, August*)	1.3	8.1	315



Figure 4 Simulation of the wave model of the east season with the coming direction of waves from the east.

3. Result

3.1. Measurement of Wave Results at The Study Site

Wave data retrieval was carried out at Panjang Island waters, Jepara, Central Java with the coordinates of 110 ° 37'59.52 "BT and 6 ° 34'44.15" LS at a depth of 11 meters. The time of data collection is carried out for five days from October 4, 2017, to October 8, 2017. The results of wave measurements are in the form of wave height and wave period, which can be seen in the following table.

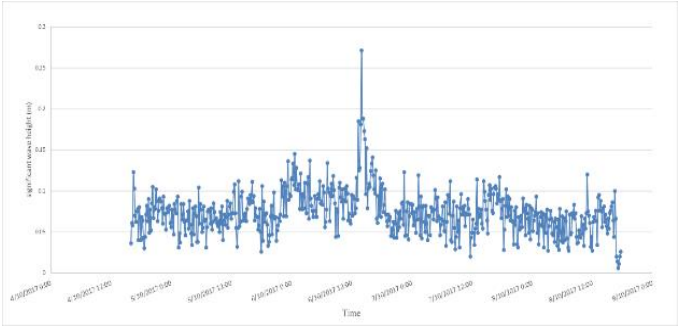


Figure. 5 Graph of significant waves high in measuring on Panjang Island Waters

Table 3. Description of significant high wave statistics measurement at Panjang Island Waters

Descriptive (Amount of Data, N = 599)		Hs (m)
Wave Height	Average	0,0700
	Minimum	0.0060
	Maximum	0.2710
	Range	0.0738
	Skewness	1.3650

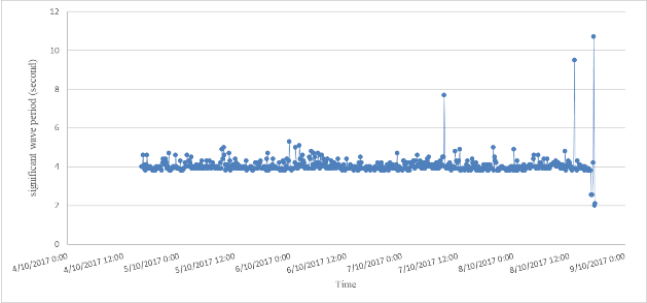


Figure. 6 Significant wave period charts measurement on Panjang Island Waters

Table 4. Descriptive statistics of significant period periods measurement at Panjang Island Waters

Descriptive (Amount of Data, N = 599)		Ts (s)
Wave Period	Average	3.567
	Minimum	2.000
	Maximum	10.70
	Range	8.700
	Skewness	2.863

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Do not use abbreviations in the title or heads unless they are unavoidable.

3.2. Simulation of Submerged Breakwater Wave Model on Panjang Island Waters, Jepara

Based on the results of calculations carried out, the waters of Panjang Island Jepara according to their relative depth, d / L of 0.105 so that these waters are included in the category of transitional waves or medium water waves. This calculation is based on the statement of Triatmodjo [11] which states that medium water waves have a relative depth between 0.05 <d / L <0.5.

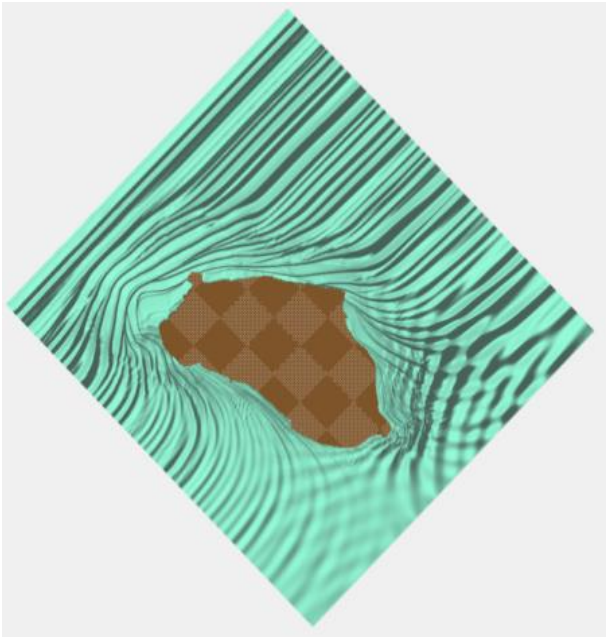


Figure 7. Propagation of Submerged Breakwater Wave Models in the West Season in Panjang Island Waters, Jepara

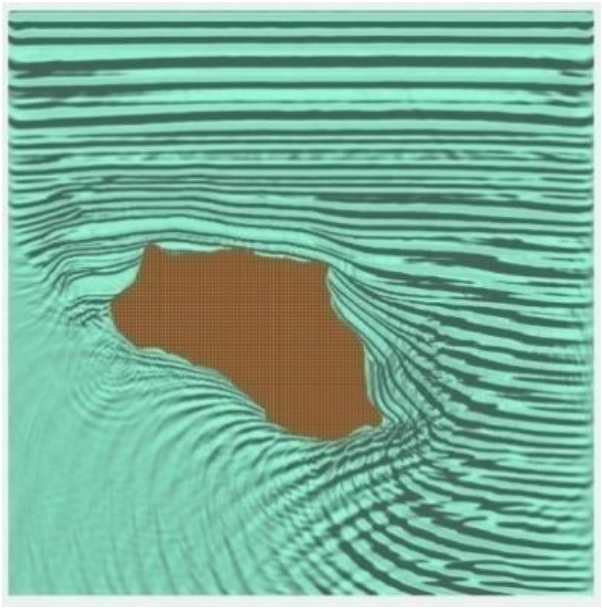


Figure 8. Propagation of Submerged Breakwater Wave Models in the first transition season in Panjang Island Waters, Jepara.

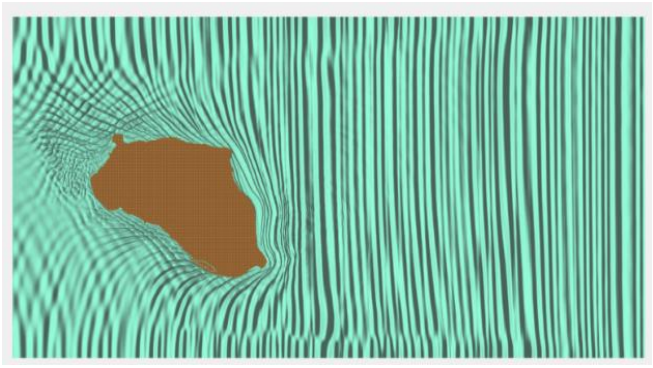


Figure 9. Propagation of Submerged Breakwater Wave Models in the East Season from The East (90°) in Panjang Island Waters, Jepara.

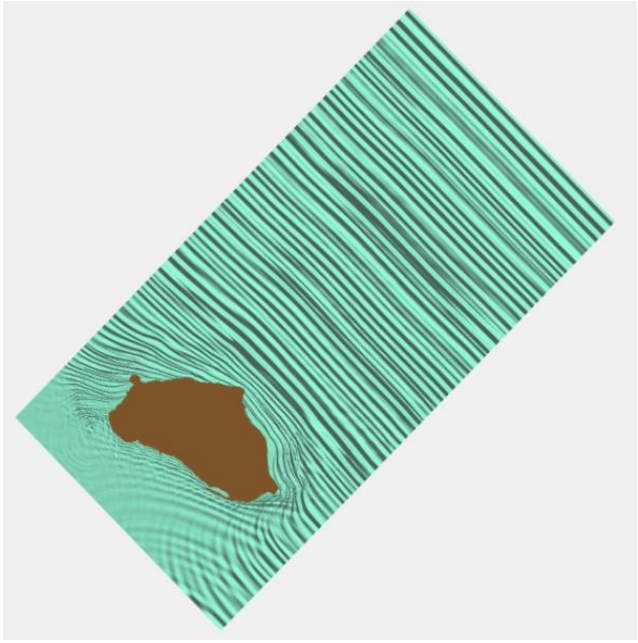


Figure 10. Propagation of Submerged Breakwater Wave Models in the East Season from the Northeast (45°) in Panjang Island Waters, Jepara.

Method to calculate the effectiveness of submerged breakwater is by comparing wave height before and after the construction of the submerged breakwater. The difference in value after compared is then made into percent.

Table 5. Submerged breakwater building efficiency in Panjang Island waters, Jepara (*December, January, February*)

Scheme	Wave Height (m)		
	Point A	Point B	Point C
Before the construction of the submerged breakwater	1.028 m	0.77 m	0.7 m
After the construction of the submerged breakwater	0.56 m	0.4 m	0.5 m
Effectiveness Value (%)	45.5	48	28.5
Average (%)		40.7	

Table 6. Submerged breakwater building efficiency in Panjang Island waters, Jepara (*March until August*)

Scheme	Wave Height (m)		
	Point A	Point B	Point C
Before the construction of the submerged breakwater	2.098 m	2.06 m	2.1 m
After the construction of the submerged breakwater	1.787 m	1.7 m	1.73 m
Effectiveness Value (%)	14.82	17.47	17.6
Average (%)		16.63	

Table 7. Submerged breakwater building efficiency in Panjang Island waters, Jepara (*September, October, December*) from the east (90°)

Scheme	Wave Height (m)		
	Point A	Point B	Point C
Before the construction of the submerged breakwater	1.52 m	1.739 m	1.65 m
After the construction of the submerged breakwater	0.83 m	0.74 m	0.76 m
Effectiveness Value (%)	45.4	57.45	53.94
Average (%)		52.26	

Based on the value of maximum wave height and wave period, it is then used to calculate the wave breakdown parameters in Panjang Island waters, Jepara. Based on the calculations that have been done, the breaking wave height is 0.54 meters with the depth of the breaking wave at 0.59 meters. The breaking wave is relatively small due to the influence of wave data collection time in October where the month is included in the Transition II season where the wind strength is relatively small.

Wind data obtained from www.ogimet.com for 11 years is displayed in the form of wind roses or windrose. It can be seen from the results that the relative velocity of the biggest winds in the western season. East and transition 1 and transition 2 are relatively small, with the dominant direction coming from the south, while the south is from the land so it is not used in wave forecasting because it cannot generate waves at sea.

Table 8. Submerged breakwater building efficiency in Panjang Island waters, Jepara (*September, October, December*) from the northeast (45°)

Scheme	Wave Height (m)		
	Point A	Point B	Point C
Before the construction of the submerged breakwater	1.74 m	1.73 m	1.65 m
After the construction of the submerged breakwater	1.01 m	0.947 m	0.9 m
Effectiveness Value (%)	41.95	45.26	45.45
Average (%)		44.22	



Figure 11. Erosion in Panjang Island, Jepara calculated by GIS data.

The location and dimensions of the submerged breakwater are based on the most vulnerable locations affected by erosion, namely on the east side of Panjang Island waters, Jepara. Then the calculation of the breaking wave plan was then carried out. According to Triatmodjo [11], a wave that moves from the deep sea with a wave height greater than H0 will break at a distance greater than Xp in front of the building. While if the wave height in the deep sea is smaller than H0 it will break in buildings with a height that is smaller than the planned wave height. Calculation of building dimensions is also based on HHWL sea level, so that the construction of a submerged breakwater is effective at the highest sea level and does not appear when the water level is lowest.

Based on the simulation results of wave propagation models on the submerged breakwater in Panjang Island waters, Jepara, it was found that submerged breakwater buildings played a role in

dampening waves in these waters, where waves were seen refraction and the wave height was reduced by about 0.7 meters.

Ocean waves are the parameters of marine hydraulics that have the most role in controlling the process of coastal dynamics and changes in the back and forth of coastlines. When the wave is large, the force and the blow energy will be large too. Broken waves are closely related to transport and distribution of sediments along the coast. The coast will experience erosion as indicated by the retreat of the coastline from its original position. This erosion process occurs when there is no balance between supply and capacity of sediment transport.

Looking at the very small efficiency in the transition season and the western season this is because the building structure is concentrated in the eastern season where the east side of the island has the worst erosion. Even though the north and west sides are the side which is directly affected by the high waves that are generated by wind from wider waters. However, the worst erosion occurred on the east side of the island. This is due to the influence of other oceanographic hydro, namely ocean currents. Ocean currents in the waters of Panjang Jepara Island according to Al Ghifari [12] is a type of tidal flow. There are unique conditions in the waters between the cape in the Bulu and Panjang Island Kelurahan, which is a narrow topography that causes a greater current velocity than the surroundings. A large mass of water moving in space which gets smaller will cause the current to move faster. Hydro-oceanographic conditions such as ocean currents have a direct influence on sediment transport around the coast and cause the worst erosion not on parts of the island facing directly to wider waters on the north and west but on the east side where there is a narrow gap between Panjang Island and Tanjung in around the Port of Kartini which causes a greater current speed than the surroundings.

So that the greatest effectiveness in the east season is 52.26%. Where buildings are effective enough to dampen incoming waves, although the damping is rather small, the reason why submerged breakwater is built is an aesthetic reason, so that the aesthetic of the beach is maintained. This is because Panjang Island, Jepara is a tourist destination in Jepara Regency as well as a place of pilgrimage for tourists. So that protection remains carried out and does not interfere with the beauty offered on Panjang Island as a tourist destination in Jepara Regency.

4. Conclusion

The characteristics of sea waves in the waters of Panjang Island, Jepara include the types of swell/sea waves, namely waves that are generated by distant winds. While based on the period, including the type of gravitational waves (period 1 sec - 30 sec). Sea waves in Panjang Island waters, Jepara according to their relative depth, d / L of 0.105 so that these waters are included in the category of transitional waves or medium water waves. Wave transformations that occur due to the influence of changes in depth with coefficient Ks of 0.989 and wave refraction process with Kr coefficient of 0.88. While the Hb breaking wave height is 0.54 m with a db breaking wave depth of 0.58 m. The effectiveness of the breakwater design in submerged breakwater buildings can reduce waves by an average of 52.26% in the east season. The most important consideration is the placement of submerged breakwater buildings on the east side of Panjang Island waters, Jepara because on the east side the worst erosion occurs. The dimensions of the submerged breakwater building are placed at a depth of 3.5 meters with a building length of 50 meters, a building width of 2 meters and a building height of 2.5 meters with a gap of 10 meters as many as four modules. This building is parallel to the depth contour on the east side of Panjang Island waters, Jepara

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

We would like thank to Oceanography Department Diponegoro University, Center for Coastal Rehabilitation and Disaster Mitigation Studies (CoRem-PKMBRP) - Center of Excellence Science and Technology (PUI), LPPM Diponegoro University and Ministry of Research Technology and Higher Education (RistekDikti) Republic of Indonesia for funding our research, with grant number 385-26/UN7.P4.3/PP/2019

References

[1] Cruz, J. 2008. Ocean Wave Energy : Current Status and Future Perspectives. Springer-Verlag, Berlin Heidelberg.
[2] Aziz, M. Furqon. 2006. Gerak Air di Laut.Jurnal Oseana, VolumeXXXI (4) : 9 – 21.
[3] Dijkstra, H.A. 2008. Dynamical Oceanography. Springer-Verlag, Berlin Heidelberg.
[4] Van der Meer, J.W., 1991, Stability and Transmission at Low Crested Structures", Delft Hydraulics Publication No. 453.
[5] Pina, G.G. and J.M. Valdes F. Alarcon, 1990, Experiments on Coastal Protection Submerged Breakwaters: A Way to Look at the Results, Proc. of the 22nd Coastal Engineering Conf., Delft, the Netherlands, pp.1592-1605.
[6] Durgappa H.R., 2008, "Coastal Protection Works", Proceedings of COPEDEC VII, Dubai, UAE.
[7] Pilarczyk, K.W. 2003, "Design of Low Crested (Submerged) Structures- an Overview", Proceedings of COPEDEC VI, Colombo, Sri Lanka.
[8] Kularatne S.R., J.W. Kamphuis, and M.A. Dabees, 2008, "Morphodynamics Around Low Crested Breakwaters – a Numerical Study", Proceedings of COPEDEC VII, Dubai, UAE.
[9] Sugianto, Denny Nugroho, Muhammad Zainuri, Alfin Darari, Suripin, Suseno Darsono, dan Nur Yuwono. 2007. Wave Height Forecasting Using Measurement Wind Speed Distribution Equation in Java Sea, Indonesia. International Journal of Civil Engineering and Technology (IJCIET), Volume 8 (9), 749-760.
[10] Holthuijsen, L. H., 2007, Wave in oceanic and coastal Waters, New York: Cambridge University Press.
[11] Triatmodjo. 1999.TeknikPantai. Beta Offset, Yogyakarta.
[12] Al Ghifari, Khalid Haidar. 2018. Studi Pola Pasang Surut Arus Musim Peralihan II Di Perairan Pulau Panjang, Jepara.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 346-354 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Project-Based Learning as a Tool of Enhancing of Entrepreneurial Attitude of Students

Dana Paľová[*,1], Martin Vejačka[1], Lukáš Kakalejčík[2]

[1]*Department of Applied Mathematics and Business Informatics, Faculty of Economics, Technical University of Košice, 04001, Slovakia*

[2] *GrowthPro, Košice, 04001 Slovakia*

ARTICLE INFO

Article history:
Received: 06 December, 2019
Accepted: 13 January, 2020
Online: 07 February, 2020

Keywords:
Entrepreneurial education
Project-Based Learning
Virtual laboratory
Europe 2020

ABSTRACT

The increase in employment and productivity is crucial for the economic growth of the European Union. Specifically, the employment of young people needs to be increased. The excessively theoretical approach to education is identified as one of the most significant problems in the educational system. Acquirement of practical experience while study, presents for the students the advantage when applying for a job or starting their own business. Self-employment or entrepreneurship might be one of the partial solutions for high youth unemployment in EU countries. The information and communication technologies transform education processes already for decades. The education in the area of entrepreneurship is also the field where these transformations occur. New educational forms and tools are being used in this area. Project-Based Learning and virtual laboratory are two examples of such new forms of education supported by information and communication technologies used in entrepreneurial education. This paper aims to present how project-based learning and virtual laboratory can support entrepreneurial education on the examples of two courses offered at our faculty. Through project-based learning, participants of these courses achieve the opportunity to implement their knowledge, skills, and experience to solve practical problems. It, all together, enhances the students' motivation and involvement in the educational process. They also practical experience establishing and managing a small virtual business within the virtual laboratory of entrepreneurial education. The paper describes the setup of these courses, their content, and evaluation by students and achieved results. The findings show that project-based learning and the virtual lab supporting entrepreneurship education via electronic tools are useful ways of entrepreneurial skills development.

1. Introduction

This paper is an extension of work originally presented in MIPRO 2019 [1]. The creation of new businesses always supported the economic growth of the economy [2]. Education in the field of entrepreneurship and enhancing the entrepreneurial attitude of students should be an important part of the curricula of universities with majors in economics and business management. However, in Slovakia, universities in the economics study plans usually prepare their students for the role of employees not for the role of entrepreneurs.

Harris [3] adduced that education often does not provide practical experience with real-world entrepreneurship. On the other hand, the time during the studies at the university could be an ideal time to start-up student's business, while they have the opportunity to practically learn, cooperate with or employ their classmates allowing them to gain work experience. At the same time, the support of the student's family in the case of business failure can be counted as one of the biggest benefits. Furthermore, a lot of successful businesses occur after multiple business failures [4]. In addition, failed enterprise brings valuable experience for involved entrepreneurs, which can be used in their future careers (both as entrepreneurs or employees). Information and communication technologies (ICT) influenced the world of business in many ways. It is currently common, that ICTs are used

[*] Dana Paľová, Faculty of Economics, Technical University of Košice, Nemcovej 32, 04001, Slovakia, +421556023268, dana.palova@tuke.sk

for setting up the businesses, for promoting and distributing products, for using the banking services or for communication with authorities. The impact of ICT was also reflected in the emergence and use of electronic commerce (or e-commerce) as a still-growing distribution and marketing channel [5].

Therefore, entrepreneurship courses set up and provided at our university aim to provide the opportunity to gain practical experience with entrepreneurship for our students.

2. Theoretical Background

2.1. Entrepreneurship and unemployment of the young generation in the context of European Union growth and sustainability

Youth unemployment is one of the key economic problems in multiple EU economies. In 2018 the youth unemployment was over 15 percent in the European Union, which is higher in comparison with other productive groups of citizens [6]. One of the possible solutions on how to reduce the youth unemployment rate might be the support of their self-employment ambitions by fostering the entrepreneurial spirit in the education process. Entrepreneurship supports personal development, encourages active participation in a community or society, and teaches a person how to enter or reenter the job market as a self-employed person or an employee.

The Europe 2020 strategy urges to promote economic growth and competitiveness via innovation [7]. Entrepreneurship in innovative fields might help to achieve its goals. The individual's ability to turn ideas into action, the activity aimed at achieving the business goals and the implementation of innovation and creativity are the main characteristics of the entrepreneurship itself [8]. New Skills Agenda for Europe [9] promotes the development of the entrepreneurial skills in the Entrepreneurship Competence Framework (Entrecomp) [10]. EntreComp defines entrepreneurship as an ability necessary in the 21st century, including social, managerial, technical, cognitive and relational skills. McCorkle et al. [11] also indicate these skills as important for entrepreneurs. Wiklund and Dean [12] include also creativity as a crucial entrepreneurial skill. Table 1 (bellow) illustrates EntreComp Framework and its main parts visualization.

Table 1: EntreComp Framework

Key Area	Competency
Ideas and Opportunities	• Spotting Opportunities • Creativity • Vision • Valuing Ideas • Ethical and Sustainable Thinking
Resources	• Motivation and Perseverance • Self Awareness and Self Efficacy • Financial and Economic Literacy • Mobilizing Others • Mobilizing Resources
Into Action	• Learning Through Experience • Working with Others • Planning and Management • Taking the Initiative • Coping with Ambiguity, Uncertainty and Risk

Source: [9]

2.2. Education towards innovative entrepreneurship

When starting a new business, entrepreneurial skills are very important. To know how to reduce business risks, to know how to use the creativity tools, to know how to learn from the market and to know-how to innovate are the key competencies in the realization of a new business idea [13]. Some of the main entrepreneurial skills for successful entrepreneurship were described by the Institute of Entrepreneurship Development [14]:

- Creating a business plan and planning in general together with the access to financing are crucial factors for business sustainability.
- Communication and networking skills to clearly express ideas and create relationships with people within and outside the organization.
- Leadership and managerial skills covering skills for the ability to plan organizational activities, coordination and motivation of human resources and developing cooperation and team-work, communicating the vision of the business and its objectives, etc.
- Digital marketing skills to boost competitiveness, innovation, and productivity. These skills include the know-how in the usage of digital tools and technologies for business growth on the market.

2.3. Project-Based Learning method (PBL) in Education of 21st Century

The current generation of students was born into a world of ubiquitous Information and Communication Technologies (ICT). Therefore, it is natural for them to use ICT in every aspect of their lives, including education. Brown [15] stated that for digital youth, ICTs are an integral part of their social life, which changes their perception of life, behavior, and education. Also, their requirements on the educational process, design and content are influenced by their attitude towards ICT. Millennials spend on average 6.5 hours using ICT for different purposes (e.g. listen to and record music; view, create and publish Internet content; play video games; watch television; talk on mobile phones and instant message) [16]. They can be described by the following attributes: group-oriented, global, technologically confident, risk-taking and optimistic. Taking into account the above-mentioned facts, the International Education Advisory Board [17] defined main requirements on new learning design as following (Table 2):

Table 2: Requirements on 21st-century learning design

Requirements	Learning design
Well-defined goals	Learning objectives are clear – the teacher must define in advance clear, realistic, reachable conditions for the successful accomplishment of the learning process. The main idea is that everybody has the same criteria for passing, it doesn't matter on the way, speed and time needed to complete the objectives.
Patience	Students may redo assignments repeatedly if necessary. It is important to give a chance to try it more times until the students are finally successful. The Millenials are prepared to learn from their mistakes and they will try to do the task again and again.
Team play	Tasks are completed, evaluated and reviewed by the groups of students. Using this way, the learners achieve the possibility to introduce themselves, to identify their strong and weak characteristics and discover their uniqueness and individuality.
Tracking	Students could see their progress at any time and any particular education process level and

Requirements	Learning design
	also as a whole. It allows them to control their path to reach the learning objectives.
Change	Students are learning time management and how to progress towards (achieving) the goal by dividing the large tasks. At the same time, they can learn how the task division moves a project along.
Immediate consequences	Teachers provide immediate/useful feedback and meaningful, formative and diagnostic assessments. It helps students to identify, what works for them and what does not in an easier way.
Personalization	Students can discover their strengths and reach the learning objectives by playing different roles.
Patterns	The learning environment makes sense for the students. By learning to identify patterns and make logical decisions during the education process the environment helps students cultivate higher-level thinking skills.

Source: [17]

The project-based learning (PBL) might be one of the methods suitable for teaching this new generation of digitally native students. The Gold Standard of project-based learning (Figure 1) introduced by [18] includes also other skills such as self-reflection, critical thinking, research and inquiry skills.

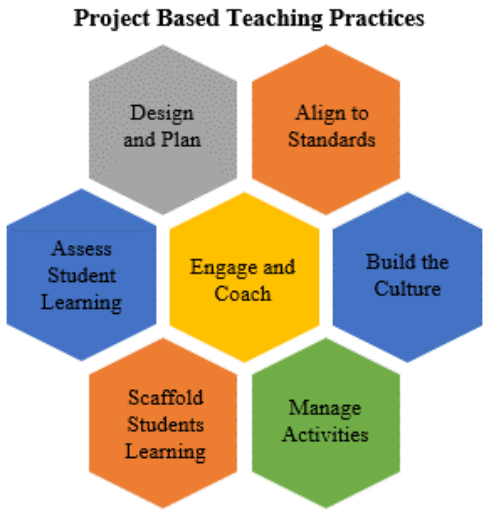


Figure 1: Key Knowledge, understanding and success skills enhanced by Gold Standard PBL; source: [18]

Project-Based Learning includes the personalization of learning for students [19]. PBL is student-directed learning that requires students to work on real-world projects over a specific period. During this time, they have to present and implement different knowledge and skills. Therefore, students find out that obtained skills and knowledge can be used in the practice [20].

While using PBL methods, lecturers could create collaborative groups of students to encourage the learning of students, as they have to face the same difficulties together [21]. The learning process then includes the utilization of the prior knowledge and information and promotes discussion in the students' groups. This leads to a higher interest of students in the educational process [22].

While using Project-Based Learning, the project solved should meet the following points:

- *Authenticity* – the project should be an authentic problem solved by using real tools, processes, and standards. The main principle is, that project needs to speak to personal concerns, interests, and issues of particular student.
- *Challenging Problem or Question* - makes learning more meaningful for students as they have to apply a lot of their knowledge and skills to find a solution to the problem.
- *Sustained Inquiry* - as students' colleagues or lecturers constantly confront their achieved results, the searching for necessary information is more dynamic and interactive. This inquiry might include interviews with experts, future customers, etc.
- *Student Voice and Choice* – students receive only project definition, but the methods of the project's working-out depend on students' choice.
- *Reflection* – supports what students have learned and encourages them to imagine how it could be applied elsewhere - outside the project. Teacher together with the students reflects the quality of student work, obstacles that arise and together try to find the problem overcoming strategies.
- *Revision* and *Critique* – show how to receive (and also give) constructive feedback during the project development (thanks to that students get experience how it runs in real life). Experts from practice or classmates might be involved in a critique.
- *Public Product* – increases motivation to higher quality work of students, because they are afraid of possible negative feedback on their results' presentation before the audience consisting of their classmates [18].

Project-Based Learning might support creativity amongst the students and teachers. Furthermore, students deepen the context of their knowledge and develop communication skills and critical thinking as they are doing an authentic project [23].

2.4. Meaning of the virtual laboratory in modern education

Currently, one of the crucial factors influencing the competitiveness and survival of the business on the marketplace is the rate of effective exploitation of ICT. The entrepreneurs must be able to keep track of a huge amount of information and available tools and to determine its possible implementation to their business area. This fact points to the need to integrate ICT into the entrepreneurial education process. The use of information and communication technologies in education may have many forms and various intensity.

The use of a virtual laboratory for entrepreneurial education is rather new. The intensive ICT involvement in teaching and learning process could be represented by the virtual laboratory. The concept of the virtual laboratory could be described in general as an environment of various cooperating ICT systems enabling the simulation of some real-world activity [24]. The usage of the virtual laboratory in entrepreneurship education might provide real-market experience for students in a very cost-effective way. Klamma et al. stated that a virtual laboratory might be very effective in ICT usage in entrepreneurial education [24].

Typically, the virtual laboratory is considered to be an interactive environment for carrying out simulated experiments.

The laboratory can be characterized as "a playground for experimentation" providing tools that can be used to manipulate objects relevant to a specific scientific domain. This concept is relevant for both, conventional and virtual laboratories. However, there is an important distinction in terminology, where a virtual laboratory is often custom-built to perform a limited set of experiments [25].

Task of design and development of the virtual laboratory for the education is highly complex. It covers the need of the set of the skills from various areas such as pedagogy, design and visualization. This process involves also the development of texts, images, learning environments and their interactions [25]. The development of the virtual laboratory and it's implementing in the teaching process, builds up the knowledge in technology, pedagogy and content knowledge [26].

Bretz et al. [27] indicated that involving virtual laboratories to the education process brings many advantages. It helps students to understand deeper science concepts, to develop their ability to identify, evaluate and solve different kinds of problems [28]. Moreover using virtual laboratories increases students' curiosity and grows up their motivation to take part in the education actively [29].

Virtual laboratories are becoming an increasingly common tool for teaching [30]. However, this is true, especially for the fields of physics, chemistry, and biology [31]. It is very often used for entrepreneurial education because it has significant potential in this area [24]. Often, the virtual laboratory in the field of entrepreneurial education is implemented differently than in the natural sciences, where it frequently represents virtual reality simulation of a lab. When used in entrepreneurial education, the virtual laboratory more likely simulates the functionality of various parts of the economy, a firm or a firm's lifecycle. The examples of simulated objects might be business registration or interaction with the public administration authorities, payment systems, market mechanisms, online marketing tools, and many others. This brings great possibilities of gaining valuable experience for the students without the necessity of great expenses spent on running a real business and involving virtual laboratory in that field supports meaningful learning by linking the previous information with a new one [32].

3. Courses supporting Entrepreneurial Education

All startups face great challenges when trying to grip on the market and up to 90% will fail to survive [33]. Da Silva et al. [34] adduced pricing issues, market validation and ignorance of customers' demands as main problems causing startup failures. Wang et al. [35] confirmed the fitting of the products on the market as an important problem of the startups adding also the functionality of the business model and issues with the acquisition of customers. Giardino et al. [36] also included targeting too-specialized markets to the list of significant issues of startups causing their failure. In consideration of these facts, the education process of the future entrepreneurs should help them to prevent such issues.

Furthermore, mastering ICT usage is currently a crucial element of entrepreneurial success. Entrepreneurial education should reflect the reality and the usage of ICTs as the most innovative and beneficial for students. McCorkle et al. [11] stated that students studying marketing and entrepreneurial education need space for creativity, what traditional education usually limits. The overview of Granitz and Pitt [37] summarized several unusual technology tools to teach entrepreneurship and marketing.

Resulting from the research in entrepreneurial and marketing education carried out, two new university courses were introduced. The courses should allow to apply students' knowledge in entrepreneurship (and in marketing as its organic element) and acquire practical skills in real conditions with the help of modern ICT tools. The above mentioned PBL method widely supported by various ICT tools was used to develop the content of the courses.

The education towards entrepreneurship is more-less covered by most of the subjects learned at the Faculty of Economics. During their studies, students are educated in theoretical fields like management, company finance, accounting, legislation, etc. This theoretical knowledge is important to achieve, but it cannot be denoted as the engine of students' curiosity and self-involvement in their study process. That is why it is really important to provide subjects that provide them more practical insight into the "real" world of the business and entrepreneurship. This is the reason why we developed innovative courses "Marketing Tools for the Presentation and Evaluation of Business Ideas" and "Economic Information Systems". In the mentioned courses, students are led to go through almost all levels of the entrepreneurial life cycle described in [38] and at the same time, they develop their skills defined in EntreComp, which are important for start-uppers and new entrepreneurs in general.

3.1. The course: Marketing tools for the presentation and evaluation of business ideas

The first course named "Marketing tools for the presentation and evaluation of business ideas" aims to concede practical education in the area of first steps in the entrepreneurial life cycle, for example, the introduction of new idea or product, and afterward company/corporate marketing. Within this course, students solve the practical problems of developing the product, fitting the product for the market, promotion, and analytics. This allows students to experience the start of the company from scratch widely supported by various ICT tools [1].



Figure 2: Final web-site of the service supporting sale of used cars; source: own based on students' web-site

The content of the course could be described as following: students should bring the business idea of product creation and then directly verify its market readiness and response. Students gain theoretical knowledge in the area of business models, key performance indicators, business objectives, web design, online marketing and its data analysis. When solving the practical problems, students obtain practical experience with social network marketing and applications such as Google Ads, Google Analytics,

Google Trends, Google Tag Manager, Google Market Finder, WordPress, and others. To achieve the final goal, i.e. to make a decision about the competitiveness and sustainability of their product or service, students need to design a web page, which is used for collecting main data for the Google Analytics tool. Figure 2 and Figure 3 represents the preview of designed websites.



Figure 3: Final web-site of the service supporting the free call of a new service provider; source: own based on students' web-site

The course has two parts: face-to-face (F2F) part, where its content is dedicated to basic project management, various business models and their Key Performance Indicators, multiple types of marketing (email, social media, content marketing), web design, marketing research, and marketing data analysis and self-study part, where the LMS Moodle is used for publishing and management of the course materials, schedules, and activities. The acquired skills and knowledge are then used by the students during work on their projects. Students form small teams (2-5 people) and their roles in the team are assigned by themselves. The student teams present their business model, product, product's fit to market and marketing strategy. The team projects are evaluated by lecturers, people from outside (operating in the field of online marketing and web-page design) and other students with a focus on business model definition and objectives, web presentation, usage of the marketing tools, gathering of the relevant data and its analysis and product viability on the market.

3.2. The course: Economic Information Systems

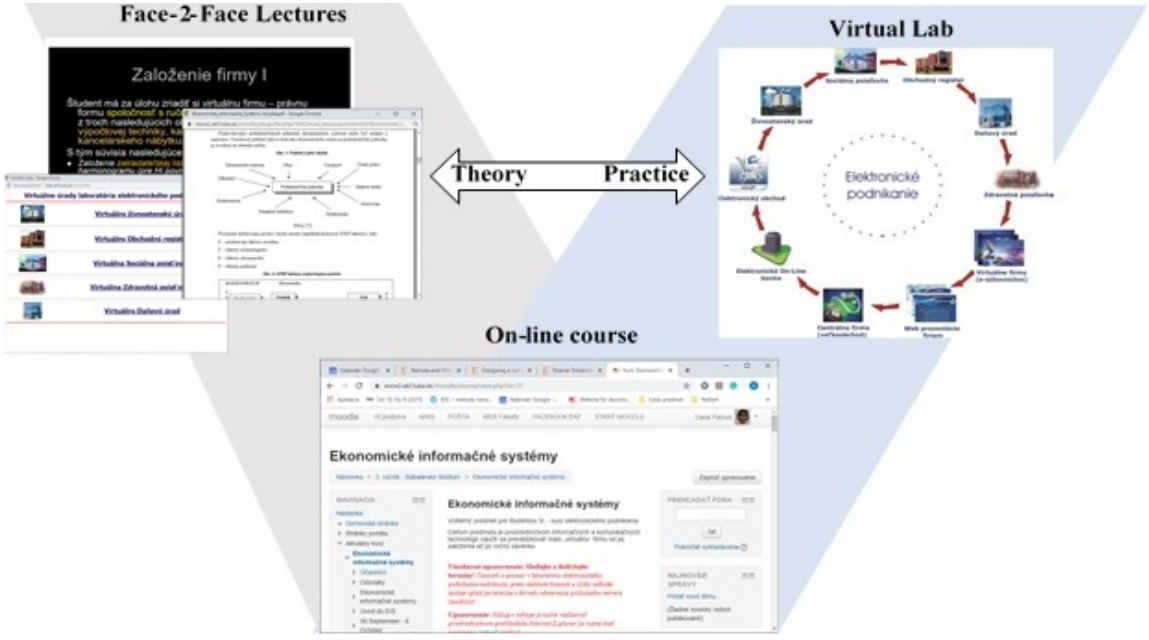
The provision of products using information and communication technologies is in the world of the business closely bounded to business intelligence systems, interchange of electronic data or electronic payments. All of these are known as electronic commerce and it can provide important tools and services for different business' relations as business-to-business, business-to-customer or customer-to-customer relations [39].

From the point of view of entrepreneurial education, it seems to be important to present the main activities needed by the entrepreneurs in the area of electronic commerce, for example, communication with the public administration authorities to meet the legal obligations, use of e-banking and e-payments [40]. The above-mentioned activities, important during the first period of the enterprise's existence, represent the content of the second developed course "Economic Information System".

The main idea of this course could be described in Figure 4.

The Virtual Lab itself is a part of e-course, which is providing students with various study materials, for example, information on how to manage a virtual firm from its establishment to the end of the first business year using ICT. The on-line course, operated by LMS Moodle, acts as the main source where all relevant information is collected. Students can find there most important dates and the description of the tasks needed to be done.

The virtual laboratory covers the following parts of electronic entrepreneurship: authorities of the business registration (providing social and health insurance registration, taxation, licensing, etc.), virtual e-shops (both retail and wholesale), virtual bank and other entities necessary during firms' life-cycle. Students establish their business after getting familiar with all the legislative requirements, therefore they need to follow the legislation of the Slovak Republic and prepare all necessary documents for establishing the business (e.g. deed of foundation, registration forms, licenses, etc.).



Afterward, they have to operate their own business using company website, e-shop. During the subject study, they act also as clients of businesses of their classmates. Using this principle, we can simulate the real market and students could achieve experience in pricing, sales and payments coordination. At the same time, students need to keep an accounting of their business and prepare the financial statements, which are bases for the final strategic study of their business. The results of the particular part of the course are evaluated by university staff – experts in particular fields, for example, accounting, legislation, information technologies, etc. To be able to provide all content of the course fully using e-way, different tools are involved within the virtual lab, such as EJBCA (the open-source PKI certification authority), Adobe Acrobat Reader DC ©, UNICOM Electronic Online Bank [5].

4. Results and discussion

The main objective of both described courses is to provide practical insight into the company's life-cycle to the students. Every course shows the different parts of the entrepreneurship. While the course "Marketing Tools for the Presentation and Evaluation of Business Ideas" covers the beginnings of the business – just from the idea appearance to a decision whether the product could survive on the marketplace, the course "Economic Information Systems" presents the process from real establishment of the company till first year anniversary. To be able to evaluate, whether our goal was achieved, both of the courses contain the evaluation of the students' skills and knowledge acquired and also the feedback, which is used to evaluate the content of the courses and way how they were provided.

Answering the electronic feedback questionnaire was last mandatory task for all course participants in both courses. Respondents (using the five-point Likert scale) should indicate the degree of agreement with the group of statements aimed at the quality and realization of particular parts of education process in the course. Gathered rates of agreement were presented in the form of percentages of agreement.

The course of "Marketing Tools for the Presentation and Evaluation of Business Ideas" is optional, what causes the number of enrolled students is usually lower than in the mandatory courses. The feedback and evaluation of students are gathered in each course iteration.

Students acquire the final evaluation of their performance at the end of the course with the maximum percentage of 100% in the total. Following table contains the results of students during the recent iteration of the course (Table 3):

Table 3: Mean values reached by students for specific parts of the projects

The part of the project evaluated:	Maximum value	Mean value
Formulation of objectives	10%	8.5%
Web design and its functionality	15%	13.5%
Web analytics	15%	14%
Content marketing	15%	11.5%
Data Analysis	10%	9.75%
Viability of product	10%	10%
Presentation of the project	25%	25%

Source: own contribution

At least one member of each team had to achieve official Google certification in Google Analytics Fundamentals and all students' teams have to achieve Google Tag Manager Fundamentals certifications. More than half of the students participating in the course successfully achieved these certificates. The teams often struggled to clearly define business goals and to set up Google Analytics, which lead to the misinterpretation or insufficiency of collected data from online channels.

Furthermore, students had to answer the feedback questionnaire at the course's end. The summary of the questionnaire's results can be following:

- 53% strongly agreed and 47% of students agreed that the content of the course met their expectations.
- 54% of responding students stated as the most interesting the part of designing the web and its further development, the marketing part of the course was indicated by13% of them and 33% did not have the favorite part of the course.
- Over 87% of students were satisfied with the educational style of the course, mainly they highlighted not obvious borderline between theoretical lectures and practical exercises. Overall, students evaluated the course as motivating and interesting.
- 87% of the respondents indicated the course as moderately difficult and 13% as very difficult.
- All students agreed that the acquired skills and knowledge during the course are applicable in entrepreneurial practice.
- All participants appreciated the necessity to use teamwork and creativity in the process of their course project development.
- All of the responding students confirmed that the completion of the course improved their understanding of the operation of the start-up.
- 93% of all students managed to assess the importance of the business idea.
- 87% of respondents indicated the improvement of their managerial and entrepreneurial skills.

During the EIS course, students can obtain a maximum evaluation of their performance at the level of 100% at the end of the course. The following table contains their results during the recent iteration of the course (Table 4):

Table 4: Mean values reached by students for specific parts of the course

The part of the course evaluated:	Maximum value	Mean value reached
Establishing a company according to legislation requirements	20%	18.13%
Business plan preparation	15%	12.58
Website preparation	10%	9.42%
Eshop administration	15%	13.55%
Accounting	20%	16.47%
Strategic study	10%	8.75%
Analysis of the information system used in the course	10%	9.12%

Source: own contribution

During each year of Electronic Information Systems course teaching, surveys to gather feedback on the course are conducted. The survey is conducted after the course completion in each curse

iteration. Investigated participants' opinions on the EIS course are divided between the parts such as starting-up the business, e-payments, communication with public administration authorities and e-commerce. The students' answers also provide suggestions for improvement of the education process in the course. The survey completion was mandatory for all participants of the course. In the most recent survey, 117 completed questionnaires were gathered.

The main results of the recently conducted survey can be summarized as follows:

- Almost 95% of students reported, that they acquired significant experience with establishing and management of the start-up company during the EIS course.
- Over 85% percent of respondents adduced gaining new information and experience with the process of the business establishment in accordance with the valid local legislation.
- Over 98% percent of participants agreed that obtaining practical experience with e-communication with virtual public administration authorities using the electronic registry was instructive.
- Almost 95% of the respondents highlighted the use of the digital signature in electronic communication with public administration authorities as valuable experience.
- Above 42% of the course participants indicated that the creation of the business plan of their virtual company was beneficial for them. All of the students already had experience with writing the business plan during previous years of their studies.
- More than 29% of students adduced that creating their company's website company was beneficial, even though they already had created a website before in other courses.
- Setting-up and managing their firm's e-shop was a significant experience for more than 78% of participants. Mainly, the operation of an e-shop from the administrator's side was highlighted by the participants.
- Almost 63% of students appreciated the possibility to try out to keep an accounting of their firms. Especially, the experience in the first business period accounting was emphasized.
- Approximately 65% of respondents also welcomed the preparation of a strategic plan of their firms for the upcoming years (beyond the horizon of the ongoing course).
- Most of the course participants (93%) evaluated the whole course of Electronic Information Systems very positively in general. Moreover, they appreciated the opportunity to practically experience the entrepreneurial processes within the virtual laboratory of electronic entrepreneurship.

In general, it can be concluded, that both described courses "Marketing Tools for the Presentation and Evaluation of Business Ideas" and "Economic Information Systems" stimulated the higher activity of the students. Furthermore, the course participants also often discussed the various issues and topics during the course lessons in order to solve the problems with projects. In spite of short time of the course duration (three months), students handled the product development, the creation of the website, to preparation and execution of marketing campaign and analysis of the data collected (as for the "Marketing Tools for the Presentation and Evaluation of Business Ideas" course) and experience the first year of existence of virtual start-up company (as for the "Economic Information Systems" course).

In both cases, students welcomed this opportunity and highly evaluated the course's benefits. The usage of ICT tools and virtual laboratory for entrepreneurial education appears to be beneficial according to our results. Also, the participants' practical experience with the projects is among the most important outcomes of both courses.

In connection with the EntreComp Framework, students achieved competencies from the parts of "Ideas & Opportunities" (especially vision, creativity, evaluation of ideas and sustainable thinking), "Resources" (financial and economic literacy, mobilizing resources) and "Into Action" (learning through experience, working with others, planning and management, coping with ambiguity, uncertainty, and risk) [10].

5. Conclusion

Despite different actions supporting the entrepreneurship from the EU like "Funding opportunities" [41], "Erasmus for Young Entrepreneurs" [42] or "EU Startup Services" [43], it is not trivial for young people to start their own business. The business sector is challenging and changing at a rapid pace and it is hard to enter this world without any previous knowledge and skills.

In [44], 34% of self-employed respondents stated the need to take an entrepreneurship course before starting up their own business, because they perceive an insufficiency of experience and skills necessary to ensure the sustainability of the business in the labor market. At the same time, from this research follows, that just less than half of EU citizens perceive that their education gave them the know-how and the skills to enable them to run a business (41%) or allowed them to understand the role of entrepreneurs in society (47%) [44].

This was the motivation to innovate courses in a way to implement most of the missing entrepreneurial skills into the education process.

The course of "Economic Information Systems" allows participants to acquire practical entrepreneurial experience in the field of e-commerce using the virtual laboratory of electronic entrepreneurship. Students gained practical experience with setting-up business according to the Slovak Republic legislation. Moreover, they experienced the usage of digital signature and the electronic registry in communication, the creation of the business plan, the preparation of the company's website with implemented e-shop and its administration. Finally, the course participants tried to keep accounting and to create a strategic plan for the future periods of their virtual firms. The use of the virtual laboratory for the purposes of entrepreneurial education is unique in educational processes of Slovak universities.

The course of "Marketing Tools for the Presentation and Evaluation of Business Ideas" allows students to experience practical education from the introduction of a new product and its marketing. Within the course, its participants solve the problems of developing the product, product's fit to the market, marketing analytics, which allows them to experience the launch of the

business supported by ICT tools. The students have to come with a business idea and confirm its viability. Furthermore, they gain theoretical and practical knowledge in business models, online marketing, web design, and data analysis. Moreover, the course participants have to design a web page, which is also used for collecting data for Google Analytics. Students also gain practical experience with network marketing, WordPress, Google Ads, Google Trends, Google Tag Manager, Google Market Finder.

In both courses, students highly evaluated the courses' benefits and the practical experience gained during the participation in the courses. By graduating described courses, participants got a chance to achieve or improve their knowledge and skills listed in EntreComp Framework. Step by step, by taking part in the time and solving the semestral projects, they built up the particular competencies needed in the business sector, such as the ability to be innovative and creative and define the problems and find suitable solutions. The usage of virtual laboratory and Project-Based Learning for entrepreneurial education appears to be beneficial according to results acquired.

Conjunction of these courses create educational synergies for better preparation of students for their possible entrepreneurial future. Passing these courses increases chances of successful self-employment of courses' graduates or even the creation of new work-positions for the others as they might start-up new businesses in real economic environment. It also might lower the unemployment rate in given economy, what is closely connected to the objectives stated in Europe 2020 strategy [7].

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This contribution was supported by the national project "Decision Support Systems and Business Intelligence within Network Economy" (Contract No. 1/0201/19) funded by Grant Agency for Science; Ministry of Education, Science, Research and Sport of the Slovak Republic.

References

[1] L. Kakalejčík, D. Paľová, "Enhancement of students' skills via project-based learning" in MIPRO 2019: 42nd international convention. Croatian Society for Information and Communication Technology, Electronics and Microelectronics, 769-774, Rijeka, Croatia, 2019. http://dx.doi.org/10.23919/MIPRO.2019.8756958

[2] J. Gerdin, G. Linton, "Contingency fit(s) in entrepreneurship research: uses and usability", Örebro University School of Business, 2016.

[3] K. Harris, Education and Knowledge: the structured misrepresentation of reality, Routledge, 2016.

[4] P. Thiel, B. Masters, Zero to One: Notes on Startups, or How to Build the Future, New York: Crown Business, 2014.

[5] M. Vejačka, "Electronic commerce entrepreneurship education using virtual laboratory" in MIPRO 2019: 42nd international convention. Croatian Society for Information and Communication Technology, Electronics and Microelectronics, 670-675, Rijeka, Croatia, 2019. http://dx.doi.org/10.23919/MIPRO.2019.8757070.

[6] Statista, "Youth unemployment rate in EU member states as of December 2018", Statista, 2018.

[7] European Commission, "EUROPE2020: A European strategy for smart, sustainable and inclusive growth", European Commission, 2010.

[8] Council of the European Union, "Council conclusions on entrepreneurship in education and training", Education, Youth, Culture and Sport Council meeting, Council of the European Union, 2014.

[9] European Commission, "Analytical Highlight Focus On Entrepreneurial skills", EU Skills Panorama, European Commission, 2015.

[10] European Commission, "EntreComp: The Entrepreneurship Competence Framework", JRC Science for Policy Report, European Commission, 2016.

[11] D. E. McCorkle, J. M. Payan, J. Reardon, N. D. Kling, "Perceptions and reality: Creativity in the marketing classroom" Journal of Marketing Education, SAGE Publications, Thousand Oaks, Canada, 29(3), 254-261, 2007. http://dx.doi.org/10.1177/0273475307306892

[12] .J. Wiklund, S. Dean, "Entrepreneurial Orientation and Small Business Performance: a Configurational Approach" Journal of Business Venturing, 20, 71-91, 2005. http://dx.doi.org/10.1016/j.jbusvent.2004.01.001

[13] .G. T. Lumpkin, G. G. Dess, "Linking two Dimensions of Entrepreneurial Orientation to Firm Performance: The Moderating Role of Environment and Industry Life Cycle" Journal of Business Venturing, 16, 429-451, 2001.

[14] Institute of Entrepreneurship Development, "4 Important Skills for Young Entrepreneurs", Institute of Entrepreneurship Development, 2017.

[15] J. S. Brown, "Learning in the Digital Age" The Internet & the University: Forum 2001, 65-91. 2001.

[16] D.F. Roberts, U.G. Foehr, V.Rideout, "Generation M: Media in the Lives of 8-18 Year-olds", A Kaiser Family Foundation Study, 2005.

[17] International Education Advisory Board, "Learning in the 21st Century: Teaching Today's Students on Their Terms", International Education Advisory Board, 2007.

[18] Buck Institute for Education, "Gold Standard PBL: Essential Project Design Elements", PBL works, Buck Institute for Education, 2018.

[19] S. Boss, "Project-Based Learning: A Short History", Edutopia, 2011.

[20] Buck Institute for Education, "What is PBL?", PBL works, Buck Institute for Education, 2019.

[21] C. E. Hmelo-Silver, "Problem-based learning: What and how do students learn?" Educational Psychology Review, 16(3), 235–266, 2004. http://dx.doi.org/10.1023/B:EDPR.0000034022.16470.f3

[22] H. G. Schmidt, S. M. M. Loyens, T. Van Gog, F. Paas, "Problem-based learning is compatible with human cognitive architecture: Commentary on Kirschner, Sweller, and Clark (2006)" Educational Psychologist 42(2), 91–97, 2007.

[23] M. J. Sousa, M. do R. Almeida, "Entrepreneurial Skills Development" in Recent Advances in Applied Economics, Lisbon, Portugal, 2014.

[24] R. Klamma, M. Spaniol, D. Renzel, "Virtual entrepreneurship lab 2.0: sharing entrepreneurial knowledge by non-linear story-telling" Journal of Universal Knowledge, 1(3), 174-198, 2006.

[25] B. S. Wästberg, T. Eriksson, G. Karlsson, M. Sunnerstam, M. Axelsson, M. Billger, "Design considerations for virtual laboratories: A comparative study of two virtual laboratories for learning about gas solubility and colour appearance" Education and Information Technologies, 24(3), 2059–2080, 2019.

[26] M.. Koehler, "What is technological pedagogical content knowledge (TPACK)?" Journal of Education, 193(3), 13-19, 2013. http://dx.doi.org/10.1177/002205741319300303

[27] S. L. Bretz, M. Fay, L .B. Bruck, M. H. Towns, "What faculty interviews reveal about meaningful learning in the undergraduate chemistry laboratory" Journal of Chemical Education, 90(3), 281-288, 2013. http://dx.doi.org/10.1021/ed300384r

[28] L. B. Bruck, M. Towns, S. L. Bretz, "Faculty perspectives of undergraduate chemistry laboratory: goals and obstacles to success" Journal of Chemical Education, 87(12), 1416-1424, 2010. http://dx.doi.org/10.1021/ed900002d

[29] A. H., Johnstone, A. Al-Shuaili, "Learning in the laboratory; some thoughts from the literature" University Chemistry Education, 5(2), 42-51, 2001.

[30] D I. Lewis, "The pedagogical benefits and pitfalls of virtual tools for teaching and learning laboratory practices in the biological sciences" The Higher Education Academy: STEM, 2014.

[31] K. Achuthan, V. K. Kolil, S. Diwakar, "Using virtual laboratories in chemistry classrooms as interactive tools towards modifying alternate conceptions in molecular symmetry" Education and Information Technologies, 23, 2499–2515, 2018. https://doi.org/10.1007/s10639-018-9727-1.

[32] A. Hakim, A. Liliasari, A. Kadarohman, Y. M. Syah, "Effects of the natural product mini project laboratory on the students' conceptual understanding" Journal of Turkish Science Education (TUSED), 13(2), 27-36, 2016.

[33] N. Patel, "90% Of Startups Fail: Here's What You Need To Know About The 10%", Forbes, 2015.

[34] F. R. da Silva, R. Fabrício, P.R. da Silva, N.V. Galegale, G. K. Kabane, "Why technology-based startups fail? An IT management approach" in Production and Operations Management Society, POMS 26th Annual Conference, Washington, D.C., USA, 2015.

[35] X. Wang, H. Edison, S. S. Bajwa, C. Giardino, P. Abrahamsson, "Key challenges in software startups across life cycle stages" in International Conference on Agile Software Development, Springer International Publishing, 169, 2016. http://dx.doi.org/10.1007/978-3-319-33515-5_14

[36] C. Giardino, S. S. Bajwa, X. Wang, P. Abrahamsson., "Key challenges in early-stage software startups" in International Conference on Agile Software Development, Springer, Cham, Switzerland, 2015. http://dx.doi.org/10.1007/978-3-319-18612-2_5

[37] N. Granitz, P. Leyland, "Teaching about marketing and teaching marketing with innovative technology: Introduction to the special edition" Journal of Marketing Education, 33(2), 127-130, 2011. http://dx.doi.org/10.1177/0273475311410844

[38] R. W. Price, Roadmap To Entrepreneurial Success Powerful Strategies for Building a High-Profit Business, Global Entrepreneurship Institute, AMACOM Books, American Management Association, 2004.

[39] E. Turban, D. King, J.K. Lee, T.P. Liang, D.C. Turban, "Electronic commerce" Prentice Hall Press Upper Saddle, NJ, USA, 2015. http://dx.doi.org/10.1007/978-3-319-10091-3

[40] V. Venkatesh, J. Y. L. Thong, F. K. Y. Chan, P. J. H. Hu, "Managing citizens' uncertainty in e-government services: the mediating and moderating roles of transparency and trust" Information Systems Research, 27(1), 1–218, 2016. http://dx.doi.org/10.1287/isre.2015.0612

[41] European Commission, "Funding opportunities for small businesses", European Commission, 2015.

[42] European Commission, "Erasmus for Young Entrepreneurs", European Commission, 2019.

[43] European commission, "EU Startup Services", European Commission, 2017.

[44] European Commission, "Entrepreneurship in the EU and beyond", Eurobarometer, European Commission, 2010.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 368-376 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Territorial Sales Redesign Using Geotechnical Tools

Diana Sánchez-Partida[*,1], Moises Ernesto Sánchez-Castro[1], José-Luis Martínez-Flores[1], Patricia Cano-Olivos[1], Martin Straka[2]

[1]*Department of Logistics and Supply Chain Management, UPAEP University, 17 Sur 901, Barrio de Santiago, CP 72410 Puebla, Puebla, México*

[2]*Faculty of Mining, Ecology, Process Control and Geotechnologies, Technical university of Kosice, Letná 9, 042 00 Košice, Slovak Republic.*

ARTICLE INFO

Article history:
Received: 16 January, 2020
Accepted: 06 March, 2020
Online: 27 March, 2020

Keywords:
Territorial Design
Geographical Information System (GIS)
Location-Allocation Problem
P-Median Problem
Constraint of impedance

ABSTRACT

In recent years there has been technological growth in the field of facility location, considering it as a significant factor to achieve greater competition among companies. The present work gives a methodological solution for a territorial redesign in a Mexican automotive parts distribution company with the support of Geotechnological tools like Geographic Information Systems (GIS) and Global Positioning System (GPS). The objective of this research is the improvement of customer service and, at the same time, a reduction in logistics costs. In this research, were evaluated and proposed two scenarios. The first the re-allocation of the current customers to the nearest distribution center, obtaining a decrease of the distance of 5.08%; and a second scenario where the opening of four peripheral warehouses reduces up to 24.6% of the total kilometers traveled. Currently, the company is evaluating which scenario is the most convenient for the company based on its resources and expansion plans. The value of this work is to be taking into account real-world results and manage a large amount of data as well as consider the constraint of impedance.

1. Introduction and Literature Review

There could be many causes and variables for which a company does not obtain or maximize their profits. Some could be market competition, a lot, or little demand for products. However, for this research, we will focus on transportation costs that the company invests in distributing the products because the charge that generates to deliver the product to the final customer, is the expenses influence to the final price of the product, consequently in the business competitiveness.

For helping to reduce transportation costs, many companies choose to open more Distribution Centers (CEDIS). Ghiani et al. [1] define a Distribution Center as "the place mainly used in the large-scale distribution for the storage of products from different manufacturers, for distribution to the points of sale (and possibly another application belonging to the firm), located in a specific geographical area." Moreover, their location must be analyzed to use as a strategy of cost reduction.

Because of opening, a CEDI is more expensive than opening a peripheral warehouse, and many companies take this option. Ghiani et al. [1] define a peripheral warehouse as "it is a warehouse served by a central warehouse and used for distribution of products to customers. Its objective is to ensure a better quality of customer service. Moreover, its operation cost is reduced because this warehouse only has an operational workforce.

There are many ways to select a correct place for the facility, but the competition and the technological innovation to this day make critical the use of different tools such as the use of Geographic Information Systems (GIS). Lo & Yeung [2] define GIS as "Computer-based systems specifically designed and implemented for two subtle but interrelated purposes, managing geospatial data, and using that data to solve spatial problems." Church & Murray [3] also define a GIS as "A system of hardware, software, and procedures designed for decision making throughout the acquisition, administration, analysis, and output of spatially referenced information."

For this study, we did an extensive search to know the applications of logistics, supply chain management, and GIS, finding a variety of applications implemented. One of the most common uses of GIS has been to support location sciences along

[*]Diana Sánchez-Partida, Email: diana.sanchez@upaep.mx

with access to necessary spatial information, including the extraction of coordinates from location points and the attribute associated with a geographical location [4, 5].

Another case is in [6] where used GIS and Global Positioning System (GPS) to track assets within the supply chain, he cites the example of Walmart that avoid increasing the inventory costs or shortages of inventory using these technologies to curb these deficiencies giving them continuous monitoring your goods. Li et al. [7] demonstrated that genetic algorithms (GA) could be used with GIS to effectively solve spatial decision problems to locate optimal places to establish facilities using data such as population and transportation, apply GIS and solve it with GA where multiple objectives are incorporated to have as exit the best locations.

The application of GIS in recent years has been extended, whether in the public or private sector. It is due to this advance that logistics and supply chain management have found in this technology like such a promising future. [8] make use of the GIS procedures together with Multicriteria Analysis (MCA) to improve the logistics for an electricity plant whose source is biomass, applied a set of criteria for the use of sustainable land. The research considers several scenarios evaluating a growing number of biomass plants and located the best location.

Another similar case is [9], where used a model with a multicriteria technique called MAIRCA that is used to classify and select suitable locations for this case, the site of an ammunition deposit. [10, 11] make use of GIS and optimization models to improve biomass availability and minimize costs and calculate delivery costs in the cotton industry. [12] propose a tool for monitoring resources by integrating a model of information construction (BIM) and GIS within a single system where the path of the chain's status is enabled, supply, and provides warning signs to ensure the delivery of materials. [13] present some other case studies where GIS and BIM are integrated, citing each platform or software used in each study.

A case of the management of the green supply chain (GSCM) is presented [14], in which it develops and implements a GIS tool on the internet to encourage the furniture industry to make a waste collection process of wood for easier recycling. The application helped identify the closest producer of a specific wood waste material in a particular geographical area so that the buyer can obtain the shortest route to collect wood waste with the minimum cost of collection.

[15] combine GIS, simulation and optimization methods for the location of biofuel facilities and design of a supply chain to minimize the cost that it causes, here the GIS were used as precursors to select locations of facilities of biofuel utilizing a series of decision factors such as candidate facilities, crop area potential, and transportation materials. This work served to identify the locations and offered support for decision-makers to determine the optimal costs, energy consumption, and emissions for the candidates.

Approaching the use of GIS applied, [16] makes use of the models of Location-Allocation, where evaluates the possibilities of these applied to the field of service geography, and it is a case study related to the spatial analysis of primary care centers in the city of Lujan, Argentina. In [17], where is analyzed data from car travel time collected through smartphones (GPS) by Google and where it applies the GIS tools and Python programming language. Establishing an initial framework as well as to extract, analyze and visualize data, as results achieve the calculation of the fluctuation during the day, the estimation of travel time variability, and an estimate of the origin-destination (OD) matrix applied to a regional analysis in the city of Kaunas City.

For last, [18] developed a desktop tool where he makes use of the Google Maps Application Programming Interface (API), where he dynamically updated the transport network and obtained a reliable estimate of the travel time matrix. The results obtained by comparing them with those obtained by running this data with the ArcGIS Network Analysis module. Thus demonstrating its advantages, one of the main criticisms compared to the ArcGIS method is that it requires prior preparation of the network and that the data is not updated continuously, unlike the Google Maps API. Also, having peak hour data, all of the above applied to a case study to estimate the travel time in the hospital accessibility assessment.

Following the last study, one of the main criticisms is the updating of spatial data. [19] evaluate the Open Street Map (OSM) data according to several quality criteria in the case developed in QGIS using existing functions and adding new sequences for the spatial data management. It allowed the researchers to evaluate the integrity of the spatial data using fundamental indicators. The study also proposed a heuristic approach to test the navigability of OSM data; the scripts developed to provide an intuitive method to evaluate OSM data based on quality indicators can be easily used to assess the suitability of data use from any region.

Another way to reduce transportation costs by [20] is to organize the sales territories for a company with 11 sales managers to be assigned to 111 sales coverage units in Mexico. The assignment problem is modeled as a mathematical program with two objective functions. One objective minimizes the maximum distance traveled by the manager, and the other objective minimizes the variation of the sales growth goals concerning the national average. A weights method is selected to solve the bi-objective non-linear mixed-integer program. Some instances are solved using commercial software with long computational times. Also, a heuristic and a metaheuristic based on simulated annealing were developed.

A well-designed territory enhances customer coverage, increases sales, fosters excellent performance and rewards systems, and lower travel costs. [21] consider a real-life case study to design the sales territory for a business sales plan. The business plan consists of assigning the optimal quantity of sellers to a territory, including the scheduling and routing plans for each seller. The problem is formulated as a combination of assignment, scheduling, and routing optimization models. The solution approach considers a meta-heuristic using a stochastic iterative projection method for large systems. Several real-life instances of different sizes were tested with stochastic data to represent rise/fall in the customers' demand as well as the appearance/loss of customers.

For this study, we created two scenarios and used geotechnical tools to evaluate the geographical distribution of current

customers, current CEDIS, and the opening of peripheral warehouses due to the company is thinking of expanding its operations. That is why two scenarios were developed to reduce the delivery distance to customers by improving the service level and thereby minimizing the logistics costs. The first proposal is the re-allocation of customers to the nearest CEDI. Moreover, the second proposal is the opening of a new peripheral warehouse that helps to improve the indicators above further and expand its operations.

2. Problem Description

Currently, the company has 11 distribution centers (CEDIS), including the headquarters in Puebla city, spread throughout the Mexican Republic which has around 710 employees, 38,000 square meters for storage in the warehouses, accounts with more than 150 national delivery routes and daily deliveries to more than 350 cities. The company was founded in 1984 when they acquired the group of retail stores called "Refacciones Puebla" (1953), to later be acquired gradually by a North American company and one of the largest business conglomerates in the world that currently operates in the United States, Canada, Mexico, and New Zealand.

The company has been undergone many changes. It started as a family business in the city of Puebla and later was expanded throughout the country. The geographical distribution of the CEDIS currently follows an ambiguous customer allocation criterion and, because they have been opened gradually and empirically, detailed location studies have not been carried out. They have opened CEDIS, where there was a higher concentration of population and vehicular fleet, thinking that this is a good market. Nevertheless, a problem of the company is that the location of the CEDIS does not serve its closest customers, causing longer delivery times, long routes, higher distribution costs, and therefore lowering the level of customer service.

Therefore, we have seen the need to give a new proposal that integrates new variables, for example, so far have not taken other criteria such as distance CEDIS to customers and equity in each one of the centers. As a result, there are no clear criteria as to what is the area of influence of each distribution center; there is no clear delimitation of the sales territories, for example, there are cases where several sales agents from other regions can coexist in the same area generating a self-internal competition.

3. Methodology

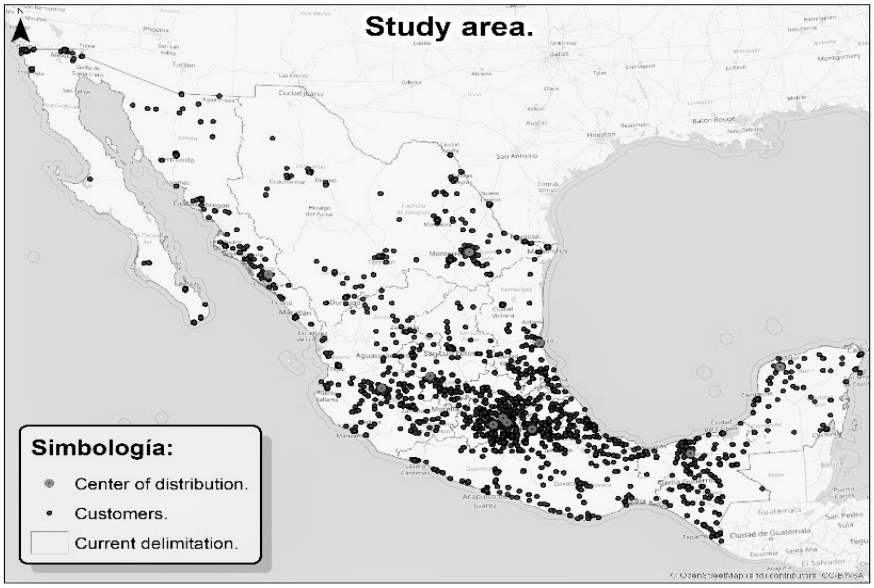
3.1. Data Collection

Burkholder [22]defines spatial data as "those numerical values which it represents, the location, size, and shape of objects found in the physical world." In our case, the spatial data corresponding to the location of the company's customers. We obtained the spatial data with the support of the field sales force of the company, approximately 140 people from the 11 CEDIS that received training.

The training consisted of the use of Geotechnological tools such as the use of the Global Positioning System (GPS) to obtain data such as Latitude, Longitude, and Business Name of the active customers of the company. The information collected in the field were validated; the points were attached to reality with the support of GIS functions to give quality to this information. For the realization of this project, several spatial data sources were used, such as internal and external data.

Internal spatial data: As mentioned at the beginning, we obtained the spatial information of the customers formed by 8911 records of customer locations, collected from July 2016 to August 2017. Moreover, we took the current geolocation of each CEDI.

External spatial data: As a source of external data, the OpenStreetMap base cartography (dataset) was used (OSMF, 2017) accessed on 10th May 2018 retrieved from (https://www.openstreetmap.org) which was a project founded in 2004 and today had positioned itself as the most famous example of Voluntary Geographical Information (VGI) on the internet [23]. The OSM project is made up of a massive global community of collaborators worldwide that contribute to keeping information on roads, trails, shops, and many more places around the world updated. We selected this cartography because it is continuously updated and useful free under an open license.



Study area. Source: self-made.

According to [24] have increased the number of developers, industries, researchers, and other users who are making use of OSM in their applications. OSM has been compared favorably with other spatial data sources regarding spatial quality, as well as citing several examples of navigation and routing applications that make direct use of OSM data.

For this work, the cartographic base of OSM was extracted to generate the road network corresponding to the update of December 2017. Then, delimit the area of influence of each distribution center with data from Municipalities of the National Institute of Statistics and Geography (INEGI, 2017) retrieved from (http://www.inegi.org.mx) accessed on 10th June 2018; corresponding to a total of 2464 municipalities that are spatially represented by polygons. All the data has been given in shapefile format (.shp) because it is considered a standard form in the GIS environment. Figure 1 shows the area of study. Currently, the company has 11 CEDIS and 8911 active customers established throughout the Mexican Republic delimited by polygons that represent the service area of each distribution center.

For this study, we developed two scenarios solved with GIS and geospatial tools based on Location-Allocation methods. In the first scenario, we evaluated how the current allocation of customers for each distribution center is compared with a reassignment of customers with the current configuration of CEDIS; that is, customers will be assigned to their nearest distribution center.

In the second scenario, we evaluated the current allocation of customers in the same way, but we proposed the opening of four peripheral warehouses following the business plan of the company. The purpose of the opening peripheral warehouses is because the company looks for improving the service level at the same time that creates new customers. A multiple regression analysis was performed to establish which municipality has better as commercial potential and where the company does not have any CEDI.

For this regression analysis, only was considered five variables because they were the most representative at the municipality level. The variables considered are the total population, automotive vehicles registered, spare parts sales business, municipal Gross Domestic Product (GDP), and personnel employed in the industrial sector. Multiple regression was run for a total of 2464 Municipalities. Then, we took the four municipalities with the most significant potential and where the company does not have a CEDI; these are Ciudad Juarez, Tijuana, Querétaro, and San Luis Potosí. Furthermore, it was generated a midpoint within the municipality to simulate the location of the peripheral warehouses. It is applied only for this study, while in the future, to have higher precision is suggested local research of facilities. The decision will be taken according to the needs of the company. Later, to identify and analyze the coverage and distances of current CEDIS to their customers, we apply a Location-Allocation model available in ArcGIS and its extension Network Analyst tools 27 [25].

The Location-Allocation model within the GIS environment has six different types of models based on various assumptions: (1) Minimize Impedance, (2) Maximize Coverage, (3) Maximize capacitated coverage, (4) Minimize Facilities, (5) Maximize Attendance, (6) Maximize Market Share y (7) Target Market Share [26]. For this study, the first model was applied due to the nature of the problem.

3.2. Minimize Impedance Problem

Minimize Impedance Problem attempts to find locations for a set of facilities that can minimize the sum of all weighted costs among all demands to reach the nearest facility. This type of problem is also known as P-Median that was developed by [27, 28]. In other words, the objective of the problem is to determine the appropriate locations for a specific number of facilities such that the total sum of the impedance weight can be minimized. That means that the demand is assigned to an installation multiplied by the impedance of the installation. The model has a set of constraints.

The first set of constraints is that any demand outside all impedance limits of the facilities will be considered unassigned. The second set of constraints is that for service installation has all the demands that fall within the impedance limit assigned to that installation. Finally, the last set of constraints is that the demands located within the impedance limit of more than one installation will be assigned to the nearest facility.

The problem of minimizing impedance has many applications in both the public and private sectors, including the location of retail stores, libraries, schools, hospitals, and others. Having as well as the main objective for the private sector to minimize costs and maximize efficiency, this model will reduce the total transportation costs; its use is the most appropriate for the private sector. Therefore, for the scenarios, it was decided to use this model to minimize the total distance.

3.3. Model Formulation of P-Median

This section shows the formulation of the P-Median problem [29]:

$$minimize \quad \sum_{i\epsilon I} \sum_{j\epsilon J} c_{ij}\, x_{ij} \tag{1}$$

$$Subject\ to \sum_{j\epsilon J} x_{ij} = 1 \ \forall i \ \in I \tag{2}$$

$$\sum_{\boldsymbol{j\epsilon J}} \boldsymbol{y_j} = \boldsymbol{p} \tag{3}$$

$$x_{ij} - y_j \leq 0 \ \forall i \ \in I; j \ \in J \tag{4}$$

$$x_{ij,}\ y_j \in \{0,1\} \forall j \in J \tag{5}$$

$$x_{ij} \geq 0 \quad \forall i \in I;\ j \in J \tag{6}$$

p The number of facilities to locate

x_{ij} The fraction of customer demand j,

which is supplied from the installation i.

$$y_i = \begin{cases} 1 \text{ If a facility is located in the candidate site } i \\ 0 \qquad\qquad\qquad\qquad \text{otherwise} \end{cases}$$

The objective function (1) minimizes the weighted total cost. Constraint (2) means that all the demand on the demand site j must be satisfied. Constraint (3) precisely requires p facilities to be located. Constraint (4) indicates that the demand nodes can only be assigned to open facilities. Constraint (5) stipulates that the location variables must be integer and binary. Finally, the constraint (6) specifies that variable assignments be nonnegative. Bear in mind that this model does not consider the constraint of impedance, it is contained in the GIS environment.

3.4. Methodology Summary

The first step is to design a network based on the updated OSM cartography. Once we had our network, to generate the scenario one, the Location-Allocation model was run in ArcGIS, taking the current customer's locations and the facilities as the current CEDIS in order to have the customers assigned to the nearest CEDI minimizing the distances traveled. Scenario two was starting from the same network. The model was again run with the 11 CEDIS plus the four peripheral ones, with a total of 15 facilities and with the same demand from current customers. Here the customers were grouped by primary areas or municipalities and were located to the territories. In Figure 2, is presented a Territorial Redesign methodology.

4. Results

As a result of the first model, which consisted of the reallocation of current customers to the nearest distribution center. We observed that there was a decrease in the total kilometers traveled between the current allocation of customers from the re-allocation result. The reduction was from 1,537,814 km to 1,459,744 km, which represents an improvement of 5.08% of the distance total traveled. Only to measure it is 7.6 times the geodesic distance from Mexico City to Moscow, that is, the shortest between two points on a curved surface, as can be seen in Table 1.

In Figure 3, it can be seen that the CEDIS of Culiacán, México Norte, Mexico Sur, León, Toluca, and Villahermosa would increase their mileage, but Puebla decreases, approximately, half of the kilometers due to the re-allocation customers to other CEDIS.

Figure 4 graphically represents the proposed re-allocation for scenario one, representing a different color on the map, which would be the service area of each current distribution center.

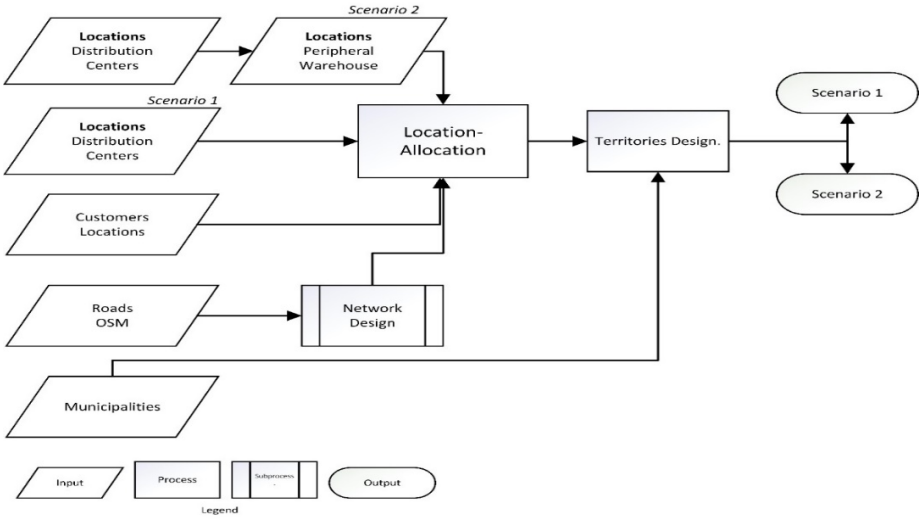


Figure 2: Proposed methodology. Source: self-made.

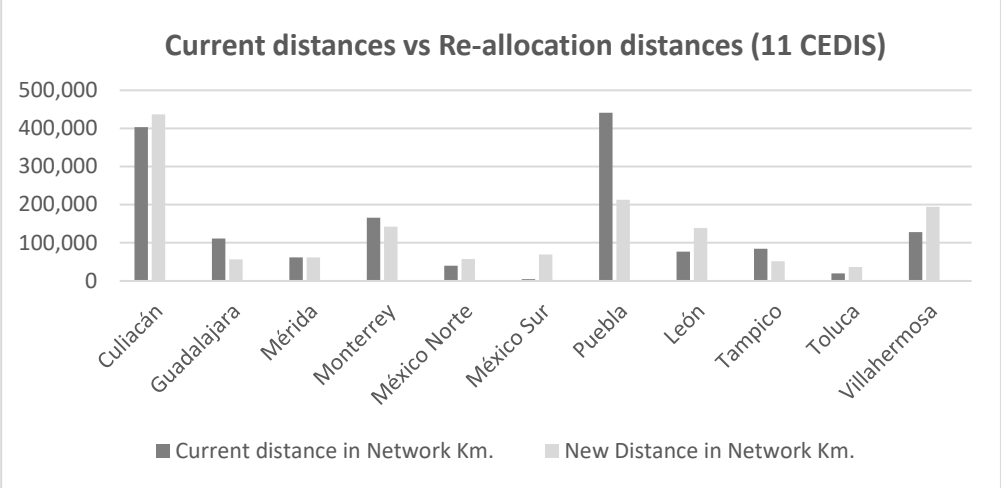


Figure 3. Current distances vs. New distances for the customers' Re-allocation (11 CEDIS). Source: self-made.

Table 1: Comparative between the current allocation and re-allocation of customers. Source: self-made.

Scenario 1						
		Current Allocation		Re-allocation		
id	Distribution center	Current distance in Network Km	Current customers	New distance in Network Km	Re-allocated customers	%
1	Culiacán	403,064	908	436,617	969	-8.32
2	Guadalajara	111,725	887	57,120	707	48.87
3	Mérida	61,684	451	61,684	450	0
4	Monterrey	165,845	940	142,056	919	14.34
5	México Norte	39,788	1,086	57,911	1110	-45.55
6	México Sur	5,042	348	69,263	717	-1273.65
7	Puebla	441,305	2,083	212,585	1445	51.83
8	León	76,691	488	138,966	709	-81.2
9	Tampico	84,619	484	52,116	406	38.41
10	Toluca	19,882	469	36,662	539	-84.4
11	Villahermosa	128,170	767	194,763	940	-51.96
	Totals	1,537,814	8,911	1,459,744	8911	5.08

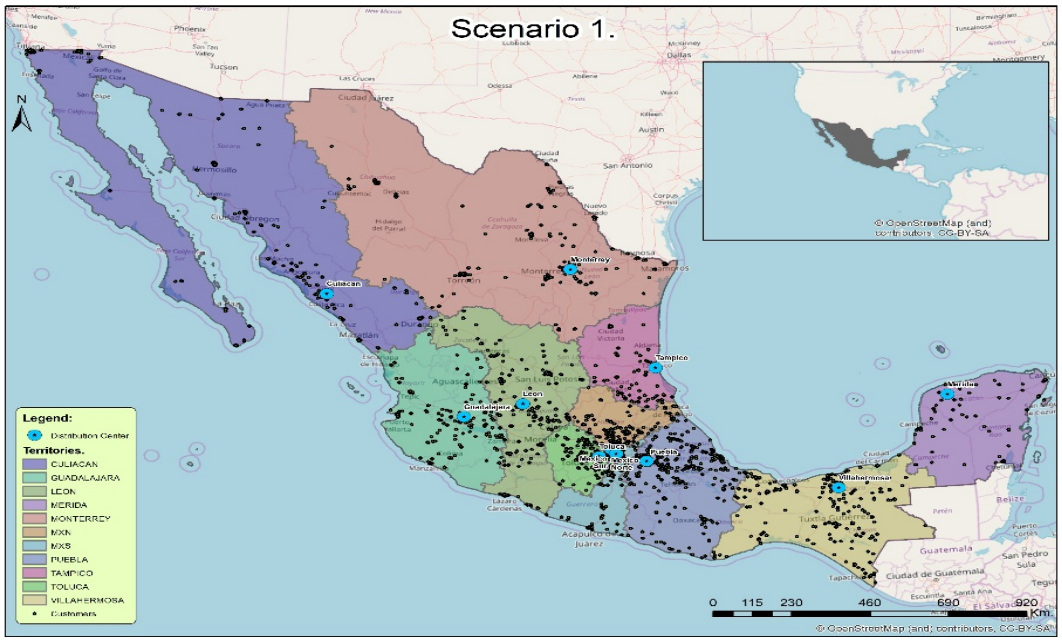


Figure 4. Territorial redesign for scenario one. Source: self-made.

For scenario two, the data was run to obtain a configuration where each current customer was re-allocated to the nearest distribution center according to the network—taking into account four peripheral warehouses that serve as auxiliary to the 11 principal CEDIS. After having solved this configuration, we obtain a reduction of 24.6% of the total trips. With this configuration, we would go from 1, 537, 814 km to 1, 159, 510 km, and there would be a reduction of 378, 304 km; it is imperative to emphasize that it is a summation of the CEDIS to each one of the customers at the national level.

With the results of this allocation, we verified that there is a significant reduction in kilometers of the CEDIS of Culiacán, Puebla, and Monterrey, which would be absorbed by Tijuana, México Sur, and Ciudad Juarez as shown in Table 2.

Table 2: Comparative results of Scenario 2. Source: self-made.

Scenario 2							
		Current Distribution			Re-allocation		
id	Distribution Center	Warehouse type	Current distance in Network Km	Current customers	New distances in Network Km	Re-allocated customers	%
1	Culiacán	Distribution center	403,064	908	92,085	736	77.1537523
2	*Tijuana*	*Peripheral warehouse*	*0*	*0*	*79,082*	*148*	-100
3	Guadalajara	Distribution center	111,725	887	54,962	698	50.8059969
4	Mérida	Distribution center	61,684	451	62,006	451	-0.5220154
5	*San Luis Potosí*	*Peripheral warehouse*	*0*	*0*	*67,329*	*391*	-100
6	Monterrey	Distribution center	165,845	940	89,886	836	45.8011999
7	Mexico Norte	Distribution center	39,788	1,086	51,710	1,098	-29.963808
8	*Ciudad Juárez*	*Peripheral warehouse*	*0*	*0*	*35,070*	*79*	-100
9	México Sur	Distribution center	5,042	348	69,282	648	-1274.0976
10	*Querétaro*	*Peripheral warehouse*	*0*	*0*	*22,375*	*220*	-100
11	Puebla	Distribution center	441,305	2,083	213,227	1,441	51.6826231
12	León	Distribution center	76,691	488	45,836	278	40.2328826
13	Tampico	Distribution center	84,619	484	47,247	389	44.165022
14	Toluca	Distribution center	19,882	469	33,721	552	-69.605673
15	Villahermosa	Distribution center	128,170	767	195,691	946	-52.680815
	Totals	15	1,537,814	8,911	1,159,510	8,911	24.6001142

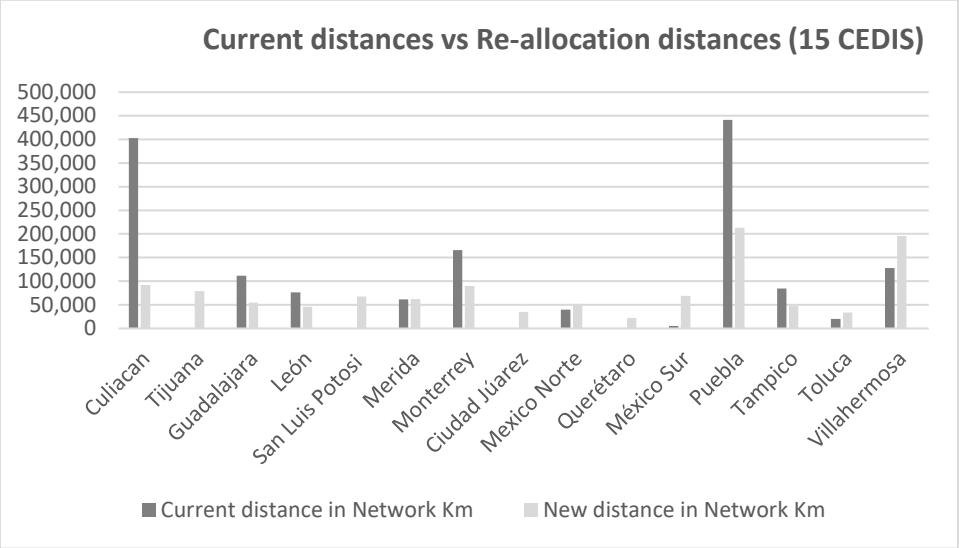


Figure 5: Current distances vs. New distances for the customers' Re-allocation (15 CEDIS). Source: self-made.

Figure 5, we can see the reductions of kilometers for Culiacan, Puebla, and Monterrey CEDIS because other CEDIS and the new proposed as peripheral warehouses like Tijuana, San Luis Potosí, Ciudad Juarez, and Querétaro now serve these customers.

As a result of the Scenario two, we can see in Figure 6, and the four other territories have been created for each new center of the peripheral distribution, as mentioned, they will be administered by another distribution center to cover the demand of the customers assigned to these territories.

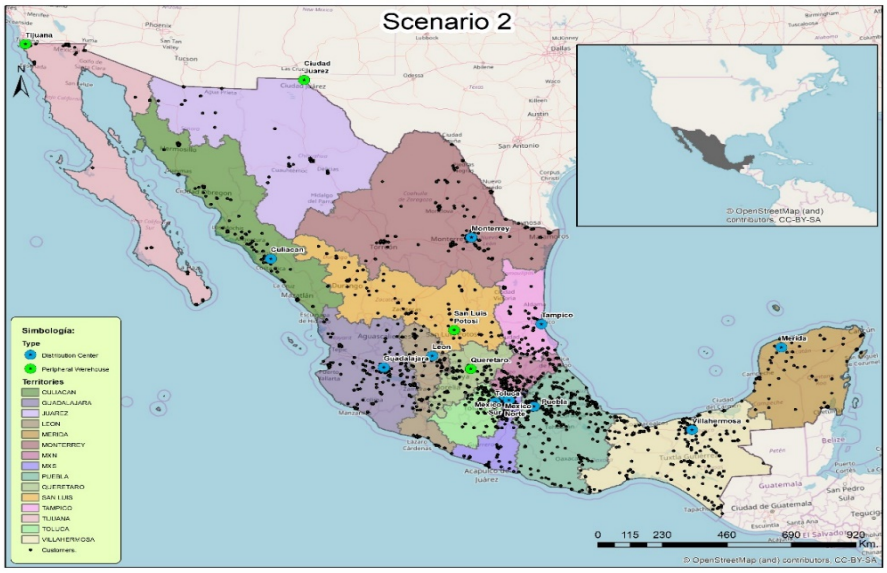


Figure 6. Territorial redesign of scenario 2. Source: self-made.

Discussion and Conclusion

Comparing the two scenarios, we can discuss that for Scenario one; there is a decrease of 5.08 % concerning the total distance. In the same way, the company will obtain a reduction in the costs of transport and raise the service level. For Scenario 2, where the 11 current CEDIS and four peripherals warehouses are considered. Likewise, there is a reduction in mileage of 24.6%. This configuration would have higher coverage to customers and lower transportation costs, but this would depend on the number of the peripheral warehouses opened. This scenario must be made an economic evaluation due to the investment that entails the opening of new facilities. The value of this work is to be taking into account real-world results and manage a large amount of data as well as consider the constraint of impedance. The continuation of this study will be compared to the operating cost of the peripheral warehouse and the variations of income for each CEDI.

In conclusion, establishing four peripheral warehouses, the routes that were considered foreign and thus more expensive because the route has a long time to arrive with the customers, now are considered as local routes. With this proposal, the local routes guarantee the delivery in a maximum 3 hours instead of 24 to 36 hours. Moreover, the proposal generates the plus of being able to attract more customers in the new territories and always improving the service level to the customers.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The corresponding author wants to acknowledge financial support from Consejo Nacional de Ciencia y Tecnología (CONACyT) with the number of scholarships (391444) and the Popular Autonomous University of the State of Puebla (UPAEP University).

References

[1] G. Ghiani, G. Laporte, R. Musmanno, Introduction to Logistics Systems Management, Wiley, 2013.
[2] C. Lo, A. Yeung, Concepts and Techniques in Geographic Information Systems, Pearson Prentice Hall, 2007.
[3] R. Church, A. Murray, Business Site Selection, Location Analysis, and GIS, Wiley, 2009.
[4] R. Church, "Location modelling and GIS." In: P. Longley, M. Goodchild, D. Maguire, D. Rhind, Geographical information systems, 293-303, Wiley, 1999.
[5] R. Church, "Geographic information systems and location science" Comput Oper Res, 29, 541–562, 2002.
[6] B. Alhenaki, "Using GIS/GPS to Optimize Supply Chain Management and Logistics at Walmart" International Journal of Scientific & Engineering Research, 7, 745-748, 2016.
[7] X. Li, A. Gar-On, "Integration of genetic algorithms and GIS for optimal location search" International Journal of Geographical Information Science, 19, 581-601, 2005.
[8] M. Delivand, A. Cammerino, P. Garofalo, M. Monteleone, "Optimal locations of bioenergy facilities, biomass spatial availability, logistics costs, and GHG (greenhouse gas) emissions: a case study on electricity productions in South Italy" Journal of Cleaner Production, 99,129-139, 2015.
[9] L. Gigovic, D. Pamucar, Z. Bajic, M. Milicevic, "The Combination of Expert Judgment and GIS-MAIRCA Analysis for the Selection of Sites for Ammunition Depots" Sustainability, 8, 1-30, 2016.
[10] F. Zhang, J. Wang, S. Liu, S. Zhang, J. Sutherland, " Integrating GIS with an optimization method for a biofuel feedstock supply chain, Biomass and Bioenergy, 98, 194-205, 2017.
[11] K. Sahoo, G. Hawkins, X. Yao, "GIS-based biomass assessment and supply logistics system for a sustainable biorefinery: A case study with cotton stalks in the Southeastern US" Applied Energy, 182, 260-273, 2016.
[12] J. Irizarry, E. Karan, F. Jalaei, "Integrating BIM and GIS to improve the visual monitoring of construction supply chain management" Automation in Construction, 31, 241–254, 2013.
[13] Z. Ma, Y. Ren, "Integrated Application of BIM and GIS: An Overview" Procedia Engineering, 196, 1072 – 1079, 2017.
[14] A. Susanty, D. Sari, W. Budiawan, Sriyanto, H. Kurniawan, "Improving green supply chain management in the furniture industry through Internet-based Geographical Information System for connecting the producer of wood waste with the buyer" Procedia Computer Science, 83, 734–741,2016.
[15] F. Zhang, D. Johnson, M. Johnson, D. Watkins, R. Froese, J. Wang, "Decision support system is integrating GIS with simulation and optimization for a biofuel supply chain" Renewable Energy, 85, 740-748, 2016.
[16] G. Buzai, "Location-allocation models applied to urban public services Spatial analysis of Primary Health Care Centers in the city of Luján, Argentina" Hungarian Geographical Bulletin, 62, 387-408, 2013.
[17] V. Dumbliauskas, V. Grigonis, A. Barauskas, "Application of Google-based data for travel time analysis: Kaunas City case study" Promet-Traffic & Transportation, 29, 613-621, 2017.

[18] F. Wang, Y. Xu, "Estimating O-D Travel Time Matrix by Google Maps API: Implementation, Advantages, and Implications" Annals of GIS, 17(4), 199-209, 2011.
[19] S. Sehra, J. Singh, R. Singh, "Assessing OpenStreetMap Data Using Intrinsic Quality Indicators: An Extension to the QGIS Processing Toolbox" Future Internet, 9, 1-22, 2017.
[20] E. Olivares-Benitez, P.N. Ibarra, S. Nucamendi-Guillén, O.G. Rojas, "Multi-Objective Territory Design for Sales Managers of a Direct Sales Company." In Optimizing Current Strategies and Applications in Industrial Engineering, 160-178, IGI Global, 2019.
[21] L. Hervert-Escobar, V. Alexandrov, "Territorial design optimization for business sales plan" Journal of Computational and Applied Mathematics, 340, 501-507, 2018, https://doi.org/10.1016/j.cam.2018.02.010
[22] E. Burkholder, "Spatial Data, Coordinate Systems, and the Science of Measurement" Journal of Surveying Engineering, 127, 143-156, 2001.
[23] J. Arsanjani, A. Zipf, P. Mooney, M. Helbich, "An introduction to OpenStreetMap in Geographic Information Science: Experiences, research, and applications." In: A. Jokar, A. Zipf, P. Mooney, M. Helbich, Eds. OpenStreetMap in GIScience, Lecture Notes in Geoinformation and Cartography, 1–18, Springer International Publishing Switzerland, 2015.
[24] P. Mooney, M. Minghini, "A Review of OpenStreetMap Data. In: Foody G. See L. Fritz S. Mooney P. Olteanu-Raimond A. Fonte C. and Antoniou V" Eds. Mapping and the Citizen Sensor, 37–59, London: Ubiquity Press, 2017.
[25] ESRI 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.
[26] ESRI 2018. Location-allocation analysis-Help | ArcGIS Desktop, Desktop.arcgis.com, Retrieved from http://desktop.arcgis.com/en/arcmap/latest/extensions/networkanalyst/locationallocation.htm#ESRI_SECTION1_F8182D9F421E4EA4AEE11E7B360E1340 (accessed on 10th May 2018).
[27] S. Hakimi, "Optimum location of switching centers and the absolute centers and medians of a graph" Oper Res, 12, 450–459, 1964.
[28] S. Hakimi, "Optimum distribution of switching centers in a communication network and some related graph-theoretic problems" Oper Res, 13, 462–475, 1965.
[29] G. Laporte, S. Nickel, F. Saldanha da Gama, Location Science, Springer, 2015.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 6, 377-385 (2019)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

A Multi-Objective Voltage Optimization Technique in Distribution Feeders with High Photovoltaic Penetration

Temitayo Olayemi Olowu, Mohamadsaleh Jafari, Arif Sarwat *

Department of Electrical and Computer Engineering, Florida International University, Miami, FL 33174

ARTICLE INFO

Article history:
Received: 29 September, 2019
Accepted: 03 November, 2019
Online: 30 December, 2019

Keywords:
Multi-objective Optimization
Smart Inverters
Voltage Control
Photovoltaic Systems
Volt-VAR Optimization

ABSTRACT

With increasing photovoltaic (PV) penetration on distribution feeders, voltage functions is always a challenge. To control these variations due to the intermittent nature of PV generation,many utility companies use the traditional voltage regulating devices such as ON/OFF load tap changers, voltage regulators, switched capacitor banks and reactors. The use of smart inverters (SI) has been reported to provide a more effective and economical way of voltage regulation on distribution feeder. It then becomes necessary to optimally control the operation of these traditional voltage regulating devices and the SIs. This paper presents a multi-objective technique that minimizes the voltage fluctuation (by implementing conservative voltage reduction), the overall system active power losses and the amount of reactive power injection from the capacitor banks with various constraints on the smart inverter reactive power injection and voltage regulator switching. The proposed algorithm is tested on an IEEE 34 node system with six units of 300kW (400kVA *SIs) PVs integrated using real data from an existing* 1.4MW *PV plant located at FIU. The Pareto optimization results of the proposed algorithm show the various optimal values of the PVs power factor, the voltage regulator settings and the capacitor reactive power injection. Using one of the Pareto optimal solutions, the results show that the system bus voltage profiles, total system power losses and capacitor reactive power injection were effectively optimized.*

1 Introduction

For environmental reasons, there has been a consistent increase in global installation of photovoltaic (PV) systems. This is further facilitated by the declining cost of PV modules over the years [1–5]. The increase is PV installations is evident as the financial investments in PV installation globally is also on the rise. The increase in levels of PV penetration have lead to several challenges which include reverse power flow, voltage control issues, power quality issues [6–9] (voltage and current harmonics, flicker, momentaries), protection coordination problems, and possible increase in system losses. [3, 10–12]. Of these issues, voltage control is one of the most important and of concern to utility companies. Traditionally, utility companies uses the legacy devices such as reactors, voltage regulators (VR), ON/OFF load tap changers (OLTC), and switched capacitor bank for voltage control [13, 14]. With the integration of intermittent renewable energy resources such as PV systems, the need for fast and more effective voltage control becomes imperative. According to IEEE 1547-2018 [15, 16], inverter-based DERs such as PVs are allowed to participate in feeder voltage regulation. According to [11, 17, 18], the use of smart inverters (SIs) could provide a more effective and economical way of voltage control especially at high levels of penetration of inverter-based DERs. Voltage optimization can be carried out with the use of several components in a distribution network. Several devices such as voltage regulators, OLTC, switched capacitor banks, reactors, energy storages and recently smart inverters (SI). SIs can be used to make significant impact on the voltage profile of the distribution system using the power factor control mode or the Volt-VAR/Watt control mode. With all these devices simultaneously present in the distribution network, there is an obvious need for optimizing their operation and also their interaction in the system. Therefore the optimal coordination of the control operations of the smart inverter and these existing legacy devices becomes necessary. This paper presents a multi-objective voltage control an optimization algorithm which is an extension of a previous work done by the authors in [19]. In order to take the advantage of the reactive power injection by the smart inverter, the reactive power injection by the switched capacitor banks are minimized as one of the optimization objectives. The smart inverters are set to constant power factor mode with the optimal value selected

*Corresponding Author: Arif Sarwat, Department of Electrical and Computer Engineering, Florida International University, Miami, FL 33174, United States, Email: asarwat@fiu.edu

between 0.6 − 1 to allow for reactive power injection. The VR and OLTC switching constraints are set within ±16. The other objective functions are to flatten the nodal voltage profile towards the lower threshold of $0.95pu$ and minimize the overall active power losses in the network. The rest of the paper is organized as follows: Section 2 presents a brief review of existing voltage control and optimization techniques; the proposed optimization algorithm and its formulation is presented in Section 3; Section 4 describes the feeder used as case study and the simulation set up; Section 5 presents the simulation results and their analyses; while Section 6 concludes the paper.

2 Voltage Control and Optimization Methods

Several voltage control techniques have been proposed and reported in literature. Active distribution networks allow for bidirectional flow of active and reactive power due to the integration of distributed renewable energy generators. An evolutionary-based algorithm for LTC and shunt capacitor Scheduling was proposed by authors of [20]. The study was done to minimize the active power loss of the system while keeping the operation of the voltage control devices and the total harmonic distortion within limits. An NREL report and publication by Ding et al [21, 22] developed an iterative optimization algorithm that coordinates the use of SI, LTCs and capacitor banks for optimal voltage regulation while implementing conservative voltage reduction technique as a means of mini zing the total system losses. The impact of the proposed optimization algorithm on the system's power quality was also analyzed. Reference [23] developed three fuzzy logic controllers that control the OLTC switching, the reactive power sharing (to enable a relaxation of the OLTC switching operation) and reduce the active power curtailment from the integrated distributed generation. Similar to [23], authors of [24] proposed a Volt-VAR optimization (VVO) algorithm that relax the voltage regulators and LTC operations as well as minimizing the amount of PV generation curtailed. The proposed VVO multistage algorithm was compared to the multi-objective VVO. Most of the previously published papers did not optimally choose the SI setting such as its power factor, minimization of reactive power injection by the shunt capacitor banks by allowing more reactive power injection by the SIs, and flattening the system's voltage profile and implementing conservation voltage reduction by pushing the nodal values to 0.95 pu. To address these research gaps, this paper proposes a multi-objective optimization algorithm that flattens the node voltage profile, minimizes the power loss in the network as well as minimizing the reactive power injection by the capacitor banks by optimally selecting the power factors of the SIs , the tap position of the voltage regulators and LTC as and the amount of reactive power to be injected by the capacitor banks.

3 Formulation of Proposed Optimization Algorithm

Objective functions

The first objective OF_1 of the proposed optimization algorithm is to flatten the voltage profile of the network as much as possible and also to make the nodal voltages close to the lower boundary of the ANSI C84.1 voltage range of $0.95pu$. The benefit of this approach is to take the advantage of conservative voltage reduction which allows the total load on the network to be reduced by simply lowering the nodal voltages.

$$OF_1 = \min \sum_{j=1}^{N_n} \left| \left((V_j^{min} - V_{j,t}) + \left| \alpha V_{j,t} - V_j^{max} \right| \right) \right| \tag{1}$$

where V_j^{min} and V_j^{max} are the minimum and maximum allowable nodal voltages respectively as expressed in equation (4), while α is a factor which is used to make the nodal voltage $V_{j,t}$ close to V_j^{min}.

Objective 2 (OF_2) is minimizes the total active power losses in the network.

$$\begin{aligned} P_{loss}^{total} &= P_{loss}^{trans} + P_{loss}^{lines} \\ OF_2 &= min \left| P_{loss}^{total} \right| \end{aligned} \tag{2}$$

where $= P_{loss}^{trans}$ is the sum of the active power loss in the transformers and P_{loss}^{lines} is the sum of the active power loss in the lines.

The third objective (OF_3) of this work is to minimize the total amount of reactive power injected by the two capacitor banks connected to buses 844 and 848.

$$OF_3 = min \sum_{cb=1}^{N_c} Q_{cb} \tag{3}$$

Where Q_{cb} is the capacitor bank reactive power injection and N_c is the total number of capacitor banks in the network.

Constraints Formulation

Bus voltage constrains is set to be within the ANSI C84.1 standards.

$$\begin{aligned} V_j^{max} \geq V_{j,t} \geq V_j^{min} \\ 1.05pu \geq V_{j,t} \geq 0.95pu \end{aligned} \tag{4}$$

Capacitor bank VAR and switching constrains is given as equation (5). The maximum reactive power injection by these capacitors are $300kVAR$ and $450kVAR$ according to the values in the feeder parameters.

$$|Q_{cbmax}^j| \geq |Q_{cb}^j| \tag{5}$$

Typical VRs and OLTCs have 32 tap steps. For each tap step, the change in voltage is typically 0.625%.

$$-16 \leq Tap_{vr,t} \leq +16 \tag{6}$$

To enable the SIs inject some reactive power into the network in order to allow for some voltage control by the SI, the PF values of the SI are set as variable for the optimization in other to get their optimal values.

$$0.6 \leq PV_{PF} \leq 1 \tag{7}$$

Table 1: PV Specifications

PVs	Maximum Power (kW)	Inverter (kVA)	Phases	Default PF
PV1 and PV5	300	400	1	0.8
PV2, PV3, PV4 and PV6	300	400	3	0.8

4 Feeder Test Case and Simulation

In order to validate the effectiveness of the proposed multi-objective optimization technique, six PVs with specifications as shown in table 1 are integrated to an IEEE 34 node distribution feeder (as shown in figure 1). For each of the PVs, a minute resolution actual global horizontal irradiance and temperature data (as shown in figures 2 and 3 respectively) are used as input to estimate the power output of the six PVs. The PV locations were selected a little farther from the substation transformer where there is a potential for some voltage regulation issues.

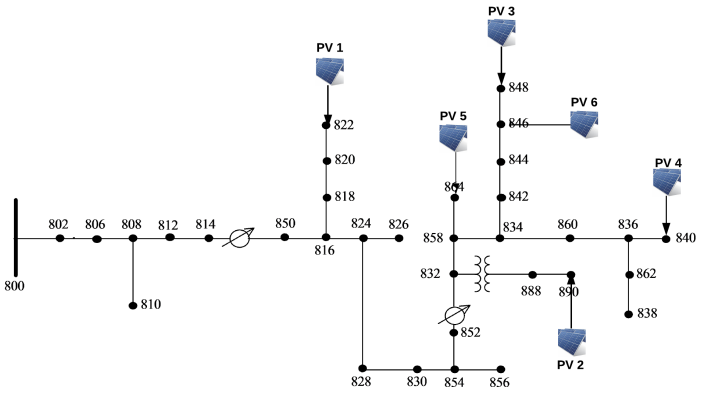


Figure 1: IEEE 34 node test feeder with PV locations

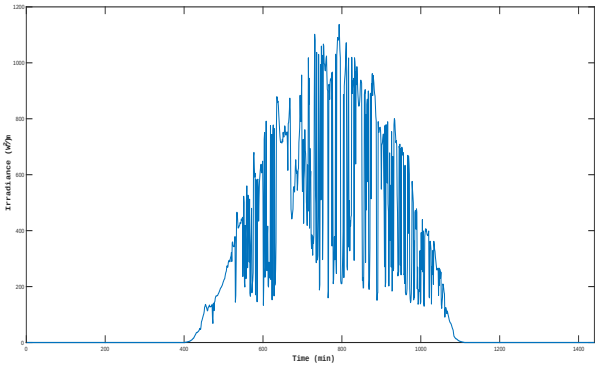


Figure 2: 1-minute resolution global horizontal irradiance profile of the location used for the simulation

5 Simulation results and Analysis

The PV integrated IEEE feeder was modelled using the OpenDSS software with the optimization algorithm developed with MATLAB. A time series simulation with a total time step of 1440 *minutes* was

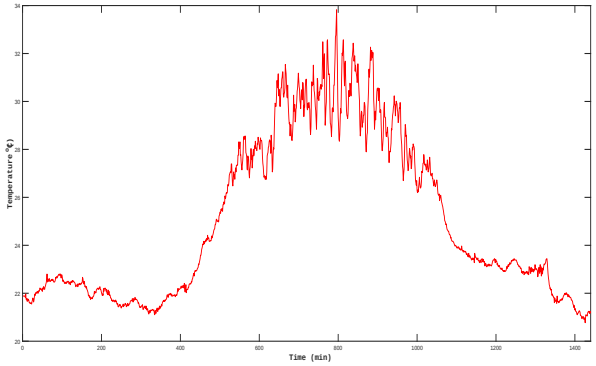


Figure 3: 1-minute resolution temperature data of the location used for the simulation

carried out. Based on the three objective functions formulated, the Pareto fronts for the 24 hour optimization are plotted as shown in figures 4, 5 and 6. The Pareto fronts show the several optimal solutions available for the multi-objective optimization.

Based on the objective of highest priority, the values of the variables which are the tap settings of the LTCs, power factor of the PVs, reactive power injection of the capacitor banks, can be selected. For example, if the objective of highest priority is to have the minimum amount of losses (OF_2) in the network the Pareto optimal values with their corresponding decision variables is as presented in tables 2 and 3. The loss minimization objective presented in tables 2 and 3 show the optimal values of the tap settings of the LTCs, power factor of the PVs, reactive power injection of the capacitor banks that will achieve the minimum amount of losses in the network. Using the optimal values acquired from the optimization algorithm, the nodal voltages of the network were simulated. The nodal voltages of buses (single phase buses) 822 and 864 are as shown in figures 9a-9b respectively while the nodal voltage for buses 840, 846, 848 and 890 are as shown figures 7, 8, 10 and 11 respectively.

The voltage profile on these buses clearly show the effectiveness of the proposed optimization algorithm. The bus voltage profiles were adequately reduced without going below the ANSI C84.1 standard voltage level of $0.95pu$. As mentioned earlier, this was done to implement the benefits of Conservative voltage reduction which can be effectively used to reduce the total load on the network as well as minimize the total system power loss. Also, as shown in figure 11, the un-optimized voltage profile of bus 890 was increased from less than the minimum $0.95pu$ to values above the minimum $0.95pu$.

Figure 12 show the un-optimized and optimized total system losses. The figure clearly shows the there is consistent reduction in the overall system losses after optimization. Figure 13 shows the capacitor reactive power with and without the use of the proposed optimization algorithm. The value of the $kVAR$ injection by the capacitor banks were consistently reduced which further shows the effectiveness of the proposed optimization algorithm.

6 Conclusions

With the consistent rise in PV penetration on distribution feeders, effective voltage regulation and optimization using Volt-VAR control becomes very important. This paper presented a multi-objective optimization using the existing network legacy devices (switched

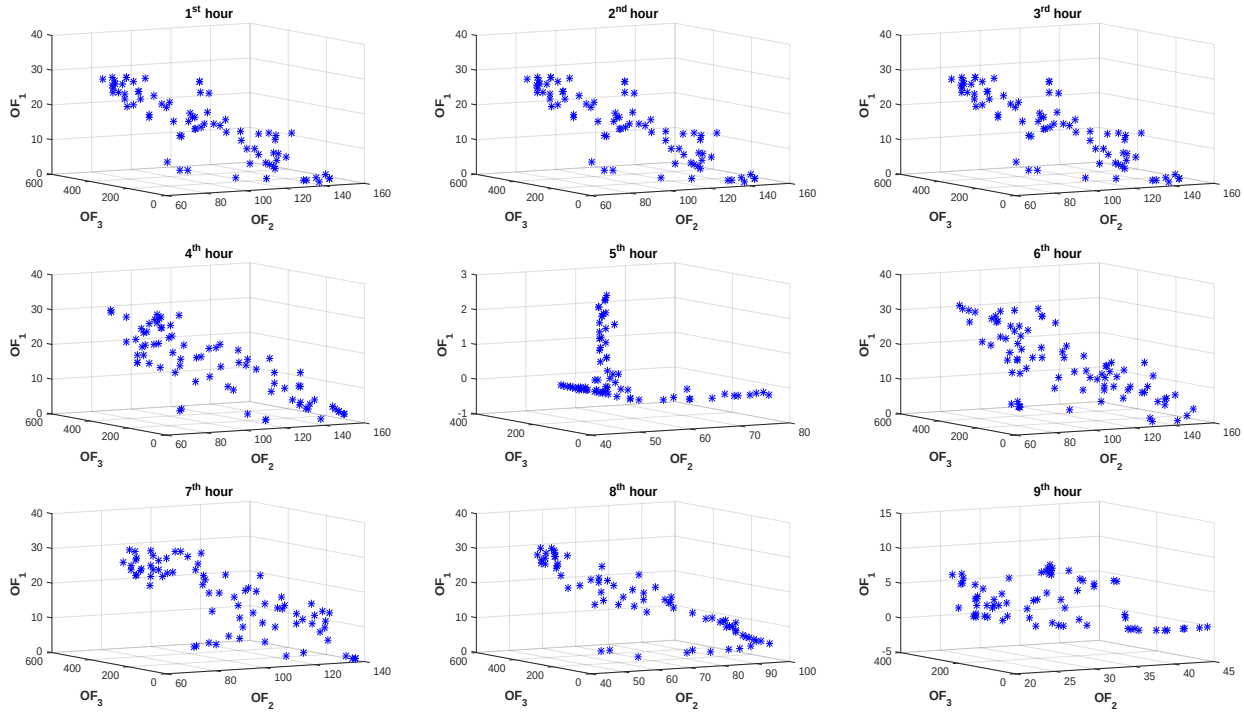


Figure 4: Hourly Pareto optimization curve for the optimized objective functions

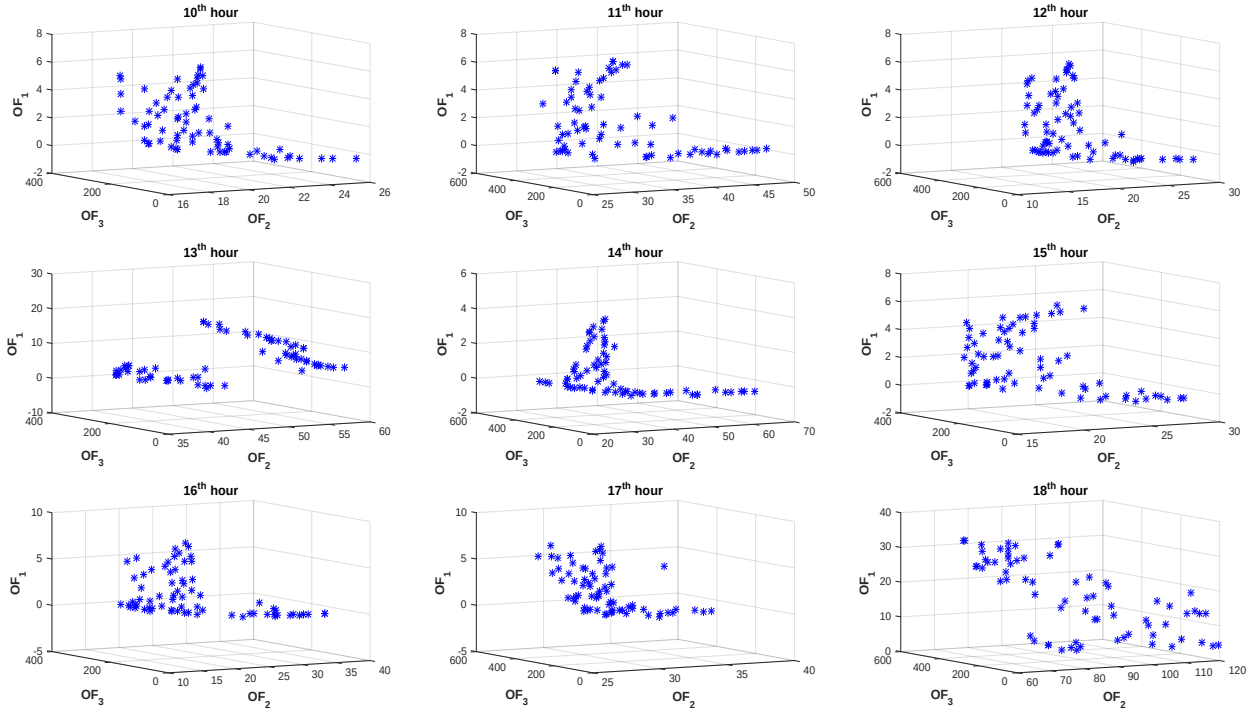


Figure 5: Hourly Pareto optimization curve for the optimized objective functions

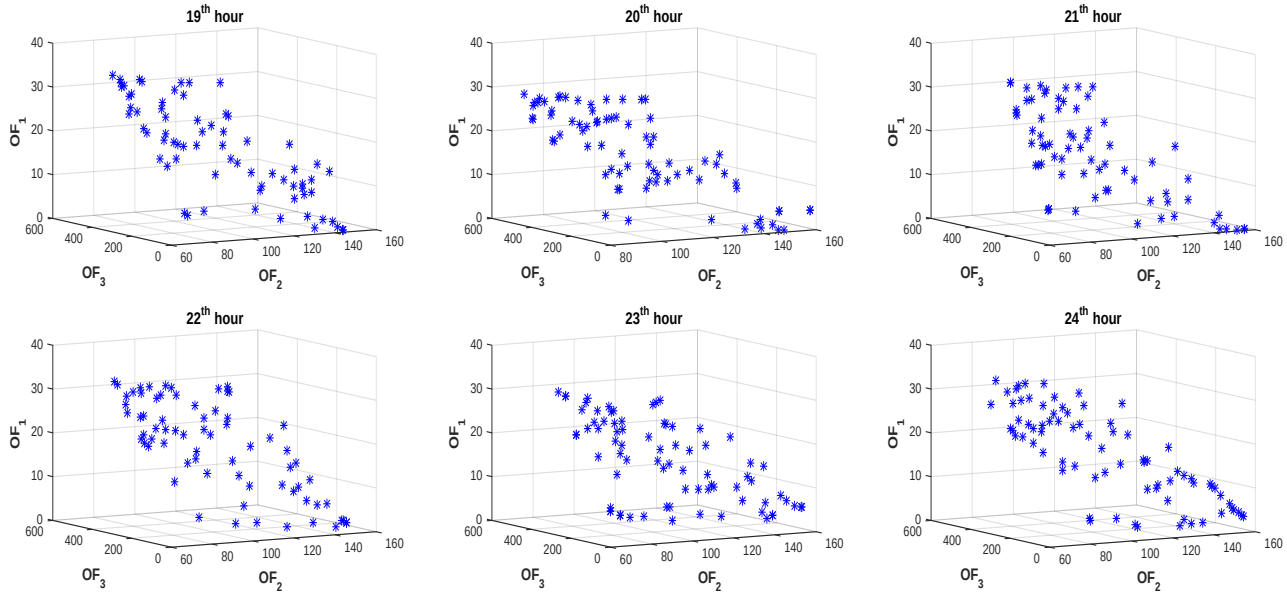


Figure 6: Hourly Pareto optimization curve for the optimized objective functions

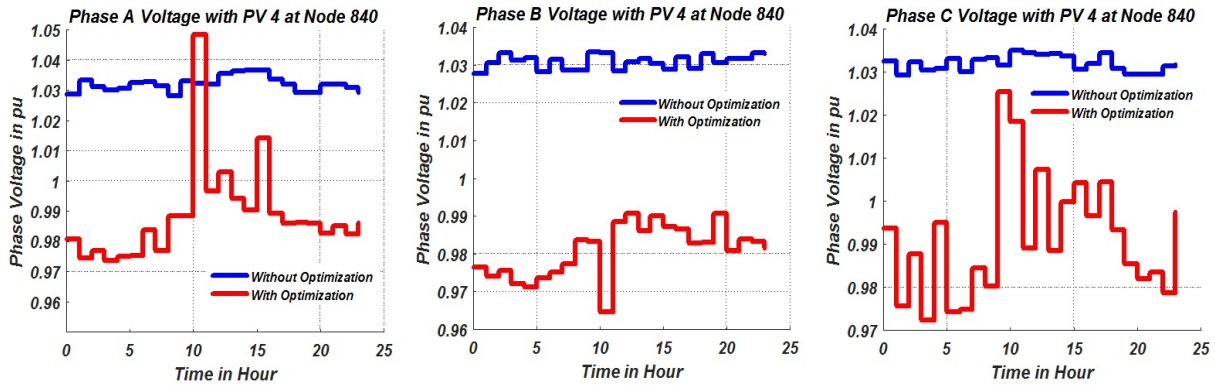


Figure 7: Voltage Profile of Node 840

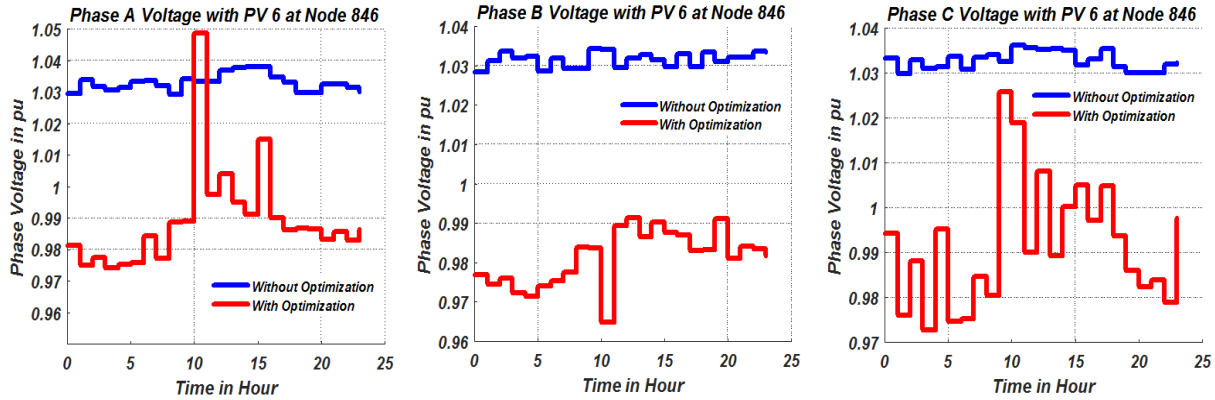
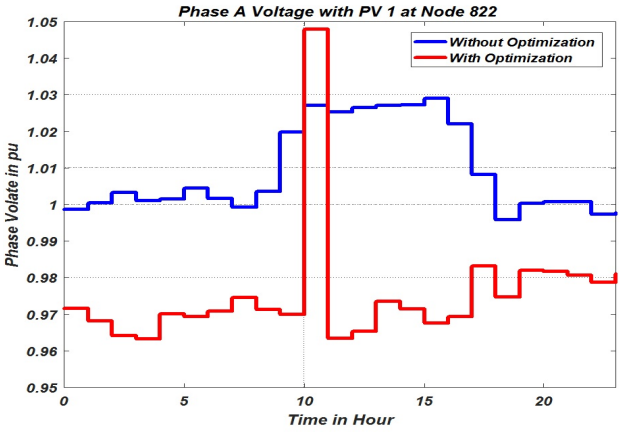
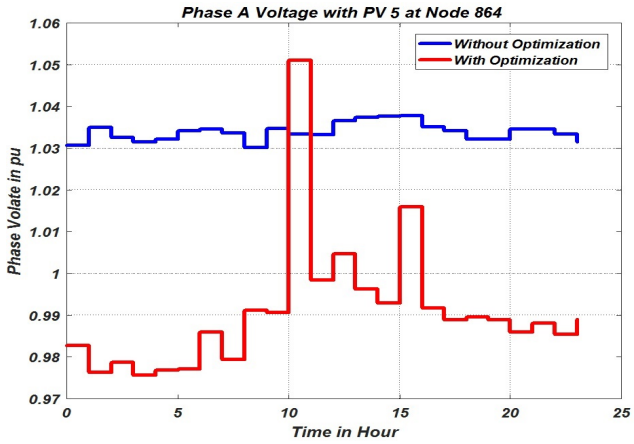


Figure 8: Voltage Profile of Node 846



(a) Voltage Profile of Node 822



(b) Voltage Profile of Node 864

Figure 9: Voltage profiles of Nodes 822 and 864

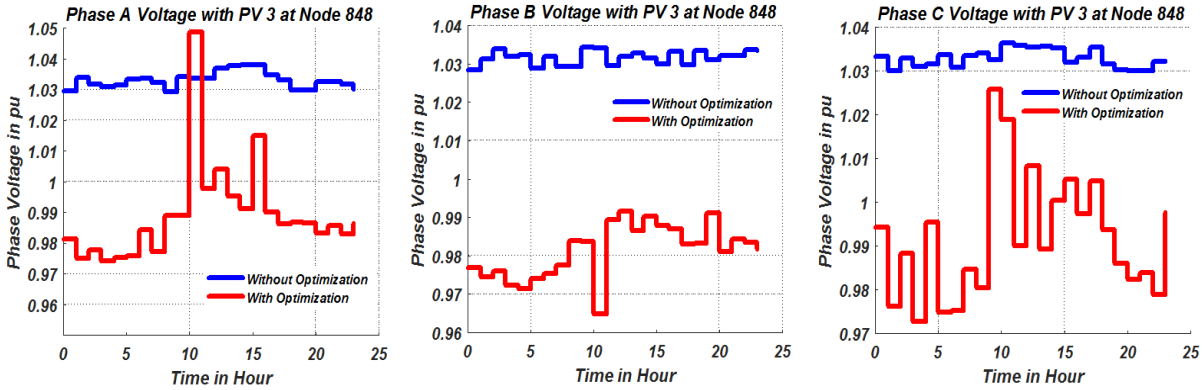


Figure 10: Voltage Profile of Node 848

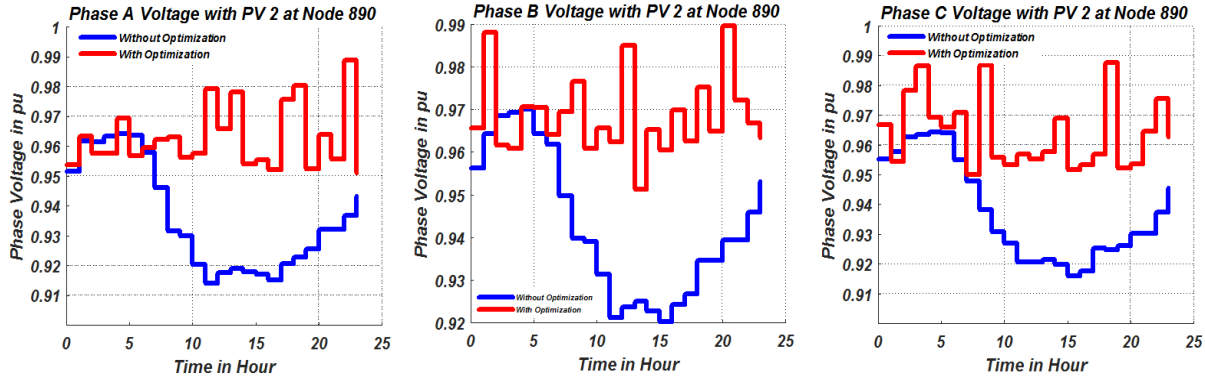


Figure 11: Voltage Profile of Node 890

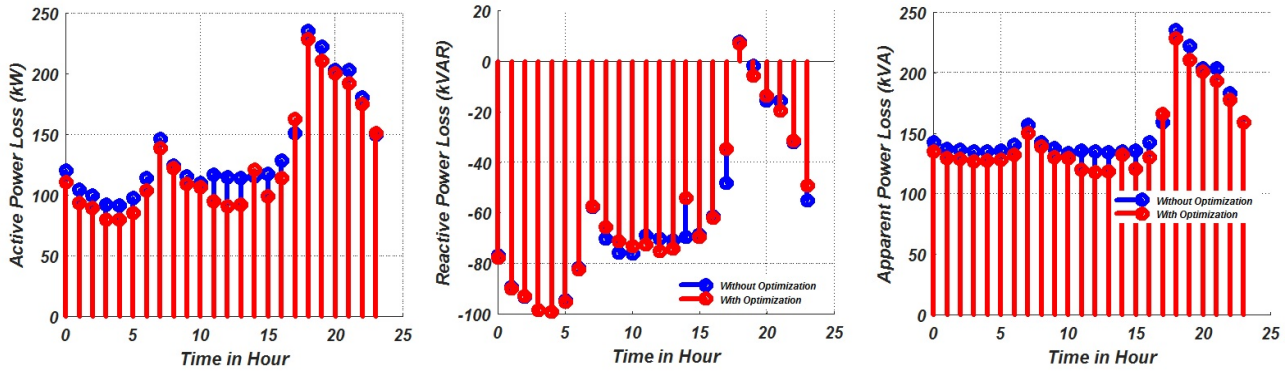


Figure 12: Hourly losses in the system

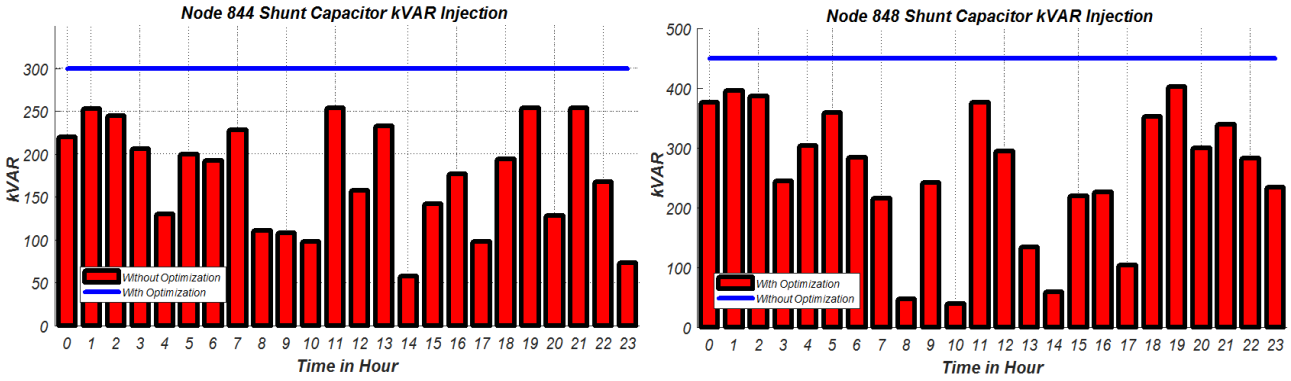


Figure 13: kVAR injection on nodes 844 and 848 by shunt capacitors

Table 2: Hourly Pareto optimal values with minimum system looses

Hour	OF_1	OF_2 (kW)	OF_3 $(kVAR)$	VR_1(pu)			VR_2(pu)		
				Ph_A	Ph_B	Ph_C	Ph_A	Ph_B	Ph_C
1	29.2225	77.8766	483.3489	0.9250	0.9500	0.9563	0.9028	0.9083	0.9093
2	29.2225	77.8766	483.3489	0.9250	0.9500	0.9563	0.9028	0.9083	0.9093
3	29.2225	77.8766	483.3489	0.9250	0.9500	0.9563	0.9028	0.9083	0.9093
4	30.5524	76.6284	458.7750	0.9250	0.9563	0.9563	0.9043	0.9162	0.9114
5	2.6370	42.8064	26.2413	1.0063	0.9625	1.000	0.9913	0.9903	0.9771
6	32.0493	75.6853	560.9548	0.9000	0.9438	0.9063	0.9406	0.9186	0.9094
7	30.4924	78.3439	540.5442	0.9125	0.9125	0.9250	0.9487	0.9086	0.9580
8	31.9572	46.2742	424.1565	0.9125	0.9188	0.9000	0.9081	0.9156	0.9031
9	6.9547	23.7373	341.3215	0.9938	1.0063	0.9625	0.9567	0.9528	0.9406
10	5.5729	17.0715	282.4728	0.9875	0.975	0.9813	1.0312	0.9834	0.9344
11	6.2144	26.3293	291.0744	0.9750	0.9938	0.9750	0.9900	0.9558	0.9468
12	5.5438	14.4402	208.2059	1.0500	1.0000	0.9750	1.0365	0.9530	0.9344
13	5.9157	38.151	275.5213	0.9813	0.9938	1.0125	0.9844	0.9778	0.9341
14	3.6769	23.0707	109.9043	1.0188	0.9563	0.9938	1.0098	0.9844	0.9509
15	5.1211	15.9318	245.9954	1.0000	1.0063	1.0063	1.0099	0.9411	0.9156
16	5.797	12.5233	228.1546	1.0688	0.9688	0.9500	1.0213	0.9782	0.9541
17	6.4897	25.0358	299.8558	0.9813	1.0250	1.0188	0.9703	0.9342	0.9021
18	33.58	66.7876	466.7909	0.9000	0.9000	0.9000	0.9342	0.9468	0.9233
19	33.5301	75.1029	562.8816	0.9000	0.9063	0.9063	0.9594	0.9022	0.9405
21	32.9089	74.7385	494.4018	0.9063	0.9250	0.9188	0.9460	0.9031	0.9085
22	32.3954	77.4375	566.4084	0.9188	0.9063	0.9	0.9578	0.9219	0.9280
23	29.8817	79.9226	576.6802	0.9063	0.9438	0.925	0.915	0.9562	0.9154
24	33.4353	73.2124	496.9962	0.9000	0.9188	0.9313	0.9275	0.9086	0.9094

Table 3: Hourly Pareto optimal values with minimum system looses

Hour	Cap_1 (kVAR)	Cap_2 (kVAR)	PV_1 (pu)	PV_2 (pu)	PV_3 (pu)	PV_4 (pu)	PV_5 (pu)	PV_6 (pu)
1	224.9997	258.3492	0.7473	0.6684	0.7505	0.8664	0.7011	0.7987
2	224.9997	258.3492	0.7473	0.6684	0.7505	0.8664	0.7011	0.7987
3	224.9997	258.3492	0.7473	0.6684	0.7505	0.8664	0.7011	0.7987
4	251.9337	206.8413	0.7418	0.8002	0.7258	0.8058	0.7102	0.8987
5	14.2472	11.9942	0.9885	0.8909	0.9125	0.9063	0.9232	1.0000
6	245.1979	315.7568	0.8264	0.8950	0.6728	0.9998	0.6260	0.9511
7	214.3476	326.1966	0.7864	0.7001	0.6404	0.8971	0.7616	0.9258
8	274.3781	149.7784	0.7423	0.6425	0.9912	0.6334	0.8202	1.000
9	186.6092	154.7122	0.9425	0.6218	0.9984	0.9275	0.9818	1.0000
10	156.2404	126.2324	0.9636	0.7052	0.9941	0.917	0.9563	1.0000
11	162.369	128.7054	0.9503	0.6230	0.9753	0.8125	0.97	1.0000
12	109.2072	98.9987	0.9600	0.7317	0.9335	0.8969	0.9441	1.0000
13	122.0062	153.5151	0.9180	0.6195	0.8228	0.8295	0.9041	0.9102
14	26.4984	83.406	0.9811	0.8656	0.9427	0.9261	0.9378	1.0000
15	158.5989	87.3964	0.9625	0.7206	0.982	0.8842	0.9549	1.0000
16	141.6355	86.5191	0.9626	0.7743	0.9764	0.907	0.9592	1.0000
17	248.5619	51.2939	0.9468	0.6161	0.9694	0.8809	0.9735	1.0000
18	258.3798	208.411	0.7905	0.7331	0.9245	0.9227	0.6792	0.9354
19	239.9177	322.9638	0.9139	0.9348	0.8009	0.6829	0.9638	0.9188
20	273.3147	292.659	0.9998	0.9995	0.9999	0.9953	0.9777	0.9625
21	66.8531	427.5486	1.0000	0.9572	0.9994	0.7732	0.8599	0.9279
22	279.084	287.3244	0.9844	0.8301	0.9429	0.9452	0.898	0.9918
23	259.5105	317.1697	0.8380	0.8885	0.7520	0.6924	0.9566	0.7904
24	231.0427	265.9535	0.8280	0.7944	0.9795	0.9783	0.8495	0.9990

capacitor banks, voltage regulators and LTCs) and respective power injection from smart inverters. The proposed algorithm minimizes the voltage fluctuation (by implementing conservative voltage reduction), the total system losses, and the amount of reactive power injection by the switched capacitor banks. The Pareto optimization results showed that by selecting the objective of highest priority, the values of the voltage regulators' tap position, reactive power injection from the capacitor banks and the power factor value of the smart inverters can be effectively selected. Using one of the Pareto optimal solutions, the bus voltage profiles, system power loss and the capacitor bank reactive power injection were simulated. The results showed the effectiveness of the proposed multi-objective algorithm.

Acknowledgment The work is a outcome of the research supported by the U.S. National Science Foundation under the grant Numbers 1553494 and 1441223.

References

[1] International Renewable Energy Agency. Data and Statistics - IRENA Resource. [Online]. Available: https://www.irena.org/en/solar

[2] T. Olowu, H. Jafari, S. Dharmasena, and A. I. Sarwat, "Photovoltaic fleet aggregation and high penetration: A feeder test case," in *SoutheastCon 2019*, April 2019, pp. 1–6.

[3] T. O. Olowu, A. Sundararajan, M. Moghaddami, and A. I. Sarwat, "Fleet aggregation of photovoltaic systems: A survey and case study," in *2019 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Feb 2019, pp. 1–5.

[4] A. Sundararajan, T. O. Olowu, L. Wei, S. Rahman, and A. I. Sarwat, "Case study on the effects of partial solar eclipse on distributed pv systems and management areas," *IET Smart Grid*, vol. 2, no. 3, pp. 477–490, 2019.

[5] F. Ariyo, B. Famutimi, T. Olowu, S. Akintade, and A. Abbas, "Web-based application for the sizing of a photovolatic (pv) solar power system," *American J. Eng. Res.(AJER)*, vol. 5, pp. 219–222, 2016.

[6] S. Rahman, M. Moghaddami, A. I. Sarwat, T. Olowu, and M. Jafaritalarposhti, "Flicker estimation associated with pv integrated distribution network," in *SoutheastCon 2018*, April 2018, pp. 1–6.

[7] A. Sundararajan, T. O. Olowu, L. Wei, S. Rahman, and A. I. Sarwat, "Case study on the effects of partial solar eclipse on distributed pv systems and management areas," *IET Smart Grid*, May 2019. [Online]. Available: https://digital-library.theiet.org/content/journals/10.1049/iet-stg.2019.0002

[8] M. Jafari, T. O. Olowu, A. I. Sarwat, and M. A. Rahman, "Study of smart grid protection challenges with high photovoltaic penetration," in *2019 North American Power Symposium (NAPS)*, Oct. 2019, pp. 1–6.

[9] A. I. Sarwat, A. Sundararajan, I. Parvez, M. Moghaddami, and A. Moghadasi, "Toward a smart city of interdependent critical infrastructure networks," in *Sustainable Interdependent Networks*. Springer, 2018, pp. 21–45.

[10] T. Olowu, A. Sundararajan, M. Moghaddami, and A. Sarwat, "Future challenges and mitigation methods for high photovoltaic penetration: A survey," *Energies*, vol. 11, no. 7, p. 1782, 2018.

[11] M. Chamana, B. H. Chowdhury, and F. Jahanbakhsh, "Distributed control of voltage regulating devices in the presence of high pv penetration to mitigate ramp-rate issues," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 1086–1095, March 2018.

[12] A. Anzalchi, A. Sundararajan, A. Moghadasi, and A. Sarwat, "High-penetration grid-tied photovoltaics: Analysis of power quality and feeder voltage profile," *IEEE Industry Applications Magazine*, vol. 25, no. 5, pp. 83–94, 2019.

[13] M. Jafari, T. O. Olowu, and A. I. Sarwat, "Optimal smart inverters volt-var curve selection with a multi-objective volt-var optimization using evolutionary algorithm approach," in *2018 North American Power Symposium (NAPS)*, Sep. 2018, pp. 1–6.

[14] T. Olowu, H. Hassan, M. Moghaddami, and A. I. Sarwat, "Physics-based design optimization of high frequency transformers for solid state transformer applications," in *2019 IEEE Industry Applications Society Annual Meeting (IAS)*. IEEE, 2019, pp. 1–6.

[15] IEEE Standards Coordinating Committee 21, "IEEE standard for interconnection and interoperability of distributed energy resources with associated electric power systems interfaces," *IEEE Std 1547-2018 (Revision of IEEE Std 1547-2003)*, pp. 1–138, April 2018.

[16] T. Olowu, S. Dharmasena, and A. I. Sarwat, "Bidirectional ac/dc converter topologies: A review," in *SoutheastCon 2019*, April 2019, pp. 1–6.

[17] F. Ding, A. Pratt, T. Bialek, F. Bell, M. McCarty, K. Atef, A. Nagarajan, and P. Gotseff, "Voltage support study of smart pv inverters on a high-photovoltaic penetration utility distribution feeder," in *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*, June 2016, pp. 1375–1380.

[18] A. Aghazadeh, M. Jafari, N. Khodabakhshi-Javinani, H. Nafisi, and H. Jabbari Namvar, "Introduction and advantage of space opposite vectors modulation utilized in dual two-level inverters with isolated dc sources," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 10, pp. 7581–7592, Oct 2019.

[19] T. O. Olowu, M. Jafari, and A. I. Sarwat, "A multi-objective optimization technique for volt-var control with high pv penetration using genetic algorithm," in *2018 North American Power Symposium (NAPS)*, Sep. 2018, pp. 1–6.

[20] A. Ulinuha, M. A. S. Masoum, S. Member, S. M. Islam, and S. Member, "Optimal Scheduling of LTC and Shunt Capacitors in Large Distorted Distribution Systems Using Evolutionary-Based Algorithms," vol. 23, no. 1, pp. 434–441, 2008.

[21] F. Ding, A. Nagarajan, S. Chakraborty, A. Nguyen, S. Walinga, M. Mccarty, F. Ding, A. Nagarajan, and S. Chakraborty, "Photovoltaic Impact Assessment of Smart Inverter Volt-VAR Control on Distribution System Conservation Voltage Reduction and Power Quality Photovoltaic Impact Assessment of Smart Inverter Volt-VAR Control on Distribution System Conservation Voltage Reduction and Power Quality," no. December, 2016.

[22] F. Ding, A. Nagarajan, M. Baggu, A. Nguyen, S. Walinga, M. Mccarty, and F. Bell, "Application of Autonomous Smart Inverter Volt-VAR Function for Voltage Reduction Energy Savings and Power Quality in Electric Distribution Systems Preprint," no. April, 2017.

[23] M. Abdelkhalek Azzouz, H. E. Farag, and E. F. El-Saadany, "Real-time fuzzy voltage regulation for distribution networks incorporating high penetration of renewable sources," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1702–1711, Sep. 2017.

[24] M. Chamana and B. H. Chowdhury, "Optimal voltage regulation of distribution networks with cascaded voltage regulators in the presence of high pv penetration," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 3, pp. 1427–1436, July 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 380-387 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Applications of Causal Modeling in Cybersecurity: An Exploratory Approach

Suchitra Abel*, Yenchih Tang, Jake Singh, Ethan Paek

Department of Computer Science and Engineering, Santa Clara University, 95053, United States

ARTICLE INFO

Article history:
Received: 16 January, 2020
Accepted: 15 May, 2020
Online: 11 June, 2020

Keywords:
Causal Modeling
Cybersecurity
Data Breach

ABSTRACT

Our research investigates the use of causal modeling and its application towards mapping out cybersecurity threat patterns. We test the strength of various methods of data breaches over its impact on the breach's discovery time as well as the number of records lost. Utilizing a Causal Modeling framework, we simulate the isolation of confounding variables while testing the robustness of varying estimators. The motivation is to shed a unique insight provided by the usage of Causal Modeling in cybersecurity.

1. Introduction

The purpose of this paper is to demonstrate the application of Causal Modeling in the domain of Cybersecurity. We engage in the scientific inquiry into the underlying causes of data breaches. Using methods of causal analysis that link concepts to observations, and a rationale connecting concepts to practice. The notion of causality, as used in Computer Science, provides principles that guide the problem specification, elaboration of the procedures, and interpretation of datasets. We employ causal modeling for the purpose of providing a computable measurement of a certain group of data breaches.

We tackle a variety of data breach problems that affect our industries and have an impact on the overall economy. In our work, we demonstrate how the usage of Causal Modeling can help us locate such data breach problems. Statistical analysis is enough for identifying associative relationships. While this is useful for general analysis, Causal Modeling provides a different structure with interventions included in it. Interventions tell us what would have happened if events other than the ones we are currently observing had happened. Such interventions allow us to avoid unnecessary steps and come directly to the point. It can also provide justification as to why and how the desired step or conclusion is arrived and provide defense for potential future cases. In order to intervene, we needed to estimate the effect of changing an input from its current value, for which no data exists. Such questions, involving estimating a counterfactual, are common in decision-making scenarios.

*Suchitra Abel, Santa Clara University Santa Clara, CA 95053, USA.
sabel@scu.edu

Statistical Prediction is the estimation of an outcome based on the observed association between a set of independent variables and a set of dependent variables. Its main application is forecasting.

Causality is the identification of the mechanisms and processes through which a certain outcome is produced. It can be used in predicting future events that are similarly connected via mechanisms and processes. Causal relations are not features that can be directly read off from the data, but have to be inferred. The field of causal discovery is concerned with this inference and the assumptions that support it.

Our research focuses on two aspects of Causal Modeling: Causal Discovery and Causal Inference. Causal Discovery algorithms try to derive causal relations from observational data. Given a set of data, a causal discovery algorithm returns a set of statements regarding the causal interactions between the measured variables.

Causal Inference is the process of drawing a conclusion about a causal connection based on the conditions of the occurrence of an effect. The main difference between causal inference and statistical inference of association is that the former analyzes the response of the effect variable when the cause is changed. The process shows causal direction, which is rarely found by statistical correlation alone. For example, a question that causal reasoning can answer is: Is there a causal link between the distribution across values of a certain variable X and values of another variable Y?

Causal inference process solves causal problems systematically, by methods such as counterfactual analysis,

graphical models, and the association between counterfactual and graphical methods.

Causal modeling resolves questions about possible causes by providing explanation of phenomena as the result of previous events. One can generate a plausible explanation for gaps within cybersecurity infrastructure. The usage of Causal Modeling can help us locate a set of data breach problems and help provide a solution for such problems.

The objective of this research is to evaluate the risks of data breach of cybersecurity incidents with the overall aim to identify patterns of importance amongst the dataset, accomplished by noting causes and effects in the modeling process. This is achieved by studying the characteristics of the VERIS Community Database (VCDB) of cybersecurity incidents. VCDB is a widely used open-source dataset containing a breadth of information regarding data breaches.

2. Background

We offer a scientific method based on the notion of causation. Following are the motivations behind the use of Causal Modeling for Cybersecurity:

One can draw from past experiences, and try to build a probability distribution [1]. Standard probability theory has been productive in these problems and similar ones, when the past experiences are readily available for analysis. But there are instances where it fails to provide adequate concepts and mathematical methods, particularly when the past experiences are either not available, or are not relevant.

A context like breach of data can interact with the phenomena of interest in ways that standard probability theory does not productively capture; that is, in ways that standard probability theory does not provide insights and methods for useful modeling and fails to capture key concepts. Some of these key concepts are the necessary and sufficient conditions that produce the essential model of the cause-effect relationships involved.

A necessary condition is one that is required if a certain effect is to follow. A sufficient condition, on the other hand, is enough for certain effects to follow.

Some of the usage of the necessary and sufficient conditions are as follows: we have to look for causes that are common in the cases where the effect also occurs. Thus, some event is not a necessary condition if it happens without the effect occurring [2,3].

We can explore causal modeling on observational data. In general, to determine whether or not an uncertain variable x_k (the supposed effect) is responsive or unresponsive to decision d we have to answer the query "Would the outcome of x_k have been the same had we chosen a different alternative for d?" Questions of this form are counterfactual queries [4,5].

We define the Counterfactual World as follows. there are some uncertain variables, X (of which x_k is an instance), such as data leakage (including some uncertainty as to why, and are we sure about the leakage?) in the scenario; there is also the set of potential causes C [6]. Possible candidates for the causes in C are:

- malware in the system
- hacking
- human error

Let U be the total set of possible effects pertaining to some scenario S. These are possibilities that should be determined correctly. There are variables $X \subseteq U$, which are uncertain variables. We also have a set of decisions D (for example, the decision that the data leakage is, indeed, there, and that it is there because of the bugs). Given these notions, the concept of counterfactual world can be defined. A counterfactual world of X and D is any instance of such world retained by $X \cup D$, after the decision maker selects a particular instance of D [7].

Definitions of unresponsiveness and responsiveness are to be understood next. Suppose that we have some uncertain variables, which form a set X. Also, suppose that we have a set of decisions D. There can be counterfactual worlds D that can form union with the set X. D is the set of scenarios where there is a list of counterfactual decisions (and the outcomes associated with the decisions), which may never take place in the real world as we encounter it. X is unresponsive to D, denoted as $X \nleftarrow D$, if X assumes the same instance in all counterfactual worlds of $X \cup D$ [8]. That is, instances of X do not affect the status of $X \cup D$. In the case of Cybersecurity, an example of a counterfactual world can be one in which no cybersecurity compromise is ever reported. These counterfactual variables are not observed, and, most probably, will never be observed. Examples of X can be concerns about Cybersecurity. These two can form a union, but X is unresponsive to D, since the instances of such concern do not affect the union. In contrast, one can think of a set X as being responsive to a set D. In this case, let the set X be the same as before, namely, the set of concerns about Cybersecurity, for example, concerns about data leakage as an element of X. The counterfactual world D can be one where Cybersecurity compromise is supposed to be reported to computer users, but ignored.

If concerns about the data leakage problem is an example of X, then it can assume different instances in different counterfactual worlds of X U D.

For example, "If one had this concern about data leakage, then one may or may not have ignored the Cybersecurity compromise report".

This shows that some instances of X can belong to some counterfactual world of X U D. Therefore, X is responsive to D.

X refers to the collections of events (indicating, for example, different states of data leakage) some of which occur after decision(s) D have been made. Given decisions D, the variables in the set C are causes of x with respect to D if all the following three conditions are met:

- Condition 1: x is not a member of C.
- Condition 2: x is responsive to D.
- Condition 3: C' is a minimal set of variables such that x is unresponsive to D in worlds limited by C' (that is, $x \leftarrow D$, and C' is a minimal set such that $x \nleftarrow c' D$).

The third condition is saying that C has a definite influence on x being responsive to D. The influence is that the relevant cause (or causes) must be included in whichever set of variables that also necessarily differ (being responsive) in accord with x being responsive to D. So, the set C' that limits the relation of x with D (regarding responsiveness) is a minimal set.

The following are the brief explanations with regard to the system discussed here.

- Condition 1 affirms that the effect (X) is not a member of the set of causes.
- Condition 2 affirms that for x (data leakage) to be caused with respect to decision D (data leakage must have been caused by the bugs in the system), it must be responsive to that decision.
- Condition 3 states the following: suppose that one can find a set of variables Y such that X, data leakage, can be different in different counterfactual worlds only when Y is different. In that case, Y must contain a set of causes.

Our approach in this paper is showing the effects of intervention. Causal modeling helps us ask the right questions about causation and helps us devise a way to emulate it by means that are not intrusive. Our emulation of interventions are based on observational studies and using data to find causal relation between them.

Causal relations are not features that can be directly read off from the data, but have to be inferred. The field of causal discovery is concerned with the inference and the assumptions that support it. Instrumental variable method ensures that we obtain the close-to-correct causal effect, even if there are unobserved conditions. Combining propensity-based and regression-based methods provides us with a causal estimate that is accurate whenever the model is correctly specified.

The potential outcomes framework can be detailed as follows: counterfactual variables such as "knowledge and action of a person P had he received the information that the cybersecurity of his computer system has been compromised "and "knowledge and action of a person P had he not received the information that the cybersecurity of his computer system has been compromised" are as appropriate as traditional variables such as "knowledge and action of a person P" – though one of these counterfactual variables is not observed, and most probably, will never be observed, in the case of this person P .

2.1. Common Cause, Confounding, Control, and Instrumental Variables.

Common causes explain the fact that there are concepts related to causation that are more important than correlation.

Suppose that a person has received a "Compromised Host" notice from some authorities, and also his computer-savvy friend (who may or may not know about the notice) has checked this person's computer and is confident that attackers have gained unauthorized access to this person's computer. Therefore, this person is worried about cybersecurity, and would like to take steps.

What are the causes of receiving such as notice? What are the causes of this computer-savvy friend being confident that attackers have gained unauthorized access to this person's computer? If there is some disaster, it could cause the "Compromised Host" notice to go out. It could also cause one's computer-savvy friend being confident that attackers have gained unauthorized access to this person's computer.

If a disaster happens, both of these are likely. This means in a data set one can find a correlation between the two.

We know there is no causal effect of receiving a "Compromised Host" notice on one's computer-savvy friend (who may or may not know about the notice) being confident that attackers have gained unauthorized access to this person's computer, or vice versa. This is the essence of "correlation does not imply causation".

When there is a common cause between two variables, then the variables will be correlated. This is part of the reasoning behind the phrase, "There is no correlation without causation".

Suppose that we are dealing with two concepts, named A and B. If neither A nor B has been definitely known to cause the other, and the two are correlated, there must be some common cause of the two. It may not be a direct cause of each of them, but it is there somewhere. This implies that we need to control for common causes if we are trying to estimate a causal effect of A on B.

Common cause variation is fluctuation caused by unknown factors resulting in a steady but random distribution of output around the average of the data.

Suppose that we take the average of the data, and do a steady but random distribution of output around the average. There will be unknown factors that will result in that distribution. This will cause a source of variation called common cause variability. This is a measure of the potential of the process – which includes how well the process can perform, if and when special cause variation is removed. Common cause variation is also called random variation, or non-controllable variation.

If we do not include hidden common causes in our model, we will estimate causal effects incorrectly. This is similar to the notion of confounders (in this particular case, some cybersecurity disaster has happened).

Confounding variables are to be understood in terms of data generating model. Pearl defines the concept of confounding as follows: Let X denote some independent variable (for example, the "Compromised system "notice), and Y some dependent variable (the person is worried and wants to take action). We might want to estimate what effect X has on Y, without regard to other potential factors; for example, if the person is, at the same time, not feeling well. We say that X and Y are confounded by some other variable Z whenever Z is a cause of both X and Y. In our case, Z is that some cybersecurity disaster has happened.

One can state that X and Y are not confounded whenever the observationally witnessed association between them is the same as the association that would be measured in a controlled experiment, with x randomized.

An equality here can be stated as P(y | do(x)) = P(y | x); this can be verified from the data generating model provided that we have all the equations and probabilities associated with the model. This is done by simulating an intervention do (X = x) and checking whether the resulting probability of Y equals the conditional probability P (y | x).

Control is a concept related to confounders. Suppose that we are attempting to assess the effectiveness of the notice being given, from population data. The data shows that prior knowledge about such incidents (Z) influences the state of mind .e.g. worry and wanting to take action (Y). In this scenario, Z confounds the relation between X (his computer-savvy friend takes the action of telling him) and Y since Z is a cause of both X and Y.

We hope to obtain an unbiased estimate P(y | do(x)) = P(y | x). In cases where only observational data are available, an unbiased estimate can only be obtained by "adjusting" for all confounding factors, which means conditioning on their various values and averaging the result.

This gives an unbiased estimate for the causal effect of X on Y. The same adjustment formula works when there are multiple confounders except, in this case, the choice of a set Z of variables that would guarantee unbiased estimates must be done with care. One can view cause-effect relationships via directed acyclic graphs; one should also link causal parameters and observed data, such as information about the subjects studied, as well estimation of the resulting parameters.

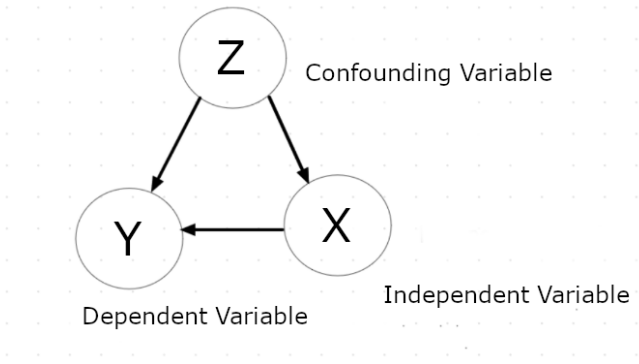


Figure 1. Example of a Causal Model

3. Overview

Since experimentation is not feasible for simulating real world data breaches, the analysis relies solely on observational data. In this regard, Judea Pearl's theory of Counterfactual World theory is extended with the use of propensity scores to calculate causal inference. The main issue to tackle regarding the use of observational data is the bias within the data caused by confounding variables, both known and unfounded. These include the previously mentioned common causes, instrumental variable, and any other covariates.

We thus present the use of causal modeling as a tool for gaining insight into how data breaches occur, and the degree to which certain associations behind these breaches can be seen as causal. We present a subset of open-sourced data offered by Verizon Communication. We then apply principles of Pearlian Causal inference through the software library DoWhy in order to understand the causal effects of our interventions.

3.1. Methodology

We concluded that DoWhy, a Microsoft open source Causal Modeling framework, was most appropriate for this current project, for its ease of use and abundant resources. It also provided an intuitive method to implement the Model -> Identify -> Estimate -> Refute structure of the analysis. All of these were readily provided by DoWhy and were thus implemented with DoWhy's built-in functions. Due to the limitation on data availability regarding data breaches, we believe these provided enough for an exploratory analysis on the subject [9,10].

DoWhy also provides a principled way of modeling a given problem as a causal graph so that all assumptions are unequivocal and explicit. It provides an integrated interface for causal inference methods, combining the two major frameworks of graphical models and potential outcomes. It also automatically tests for the validity of assumptions if possible and assesses the robustness of the estimate to violations.

It is important to note that DoWhy builds on two of the most powerful frameworks for causal inference: graphical models and potential outcomes. It uses graph-based criteria and do-calculus for modeling assumptions and identifying a non-parametric causal effect. For estimation, it switches to methods based primarily on potential outcomes.

In the following paragraphs we will describe the techniques to use for our analysis: Propensity Score Matching, Propensity Score Stratification, and Linear Regression Estimator. These techniques can all be founded within the DoWhy framework.

Linear Regression Estimator provides a baseline analysis assuming an evenly distributed dataset. It provides a foundation to compare results with the other methods. As linear regression only describes a correlation between the treatment and outcome, Propensity Score Matching and Propensity Score Stratification both use linear regression while adding additional processes in order to account for confounding variables and properly compartmentalize each data entry to find a causal relationship between the treatment and outcome.

Propensity Score Stratification takes the propensity scores of each entry and classifies them into equal sub-groups. These subgroups are classified by the similarity of the covariates. The aim is to have each sub-group represent a distribution that accurately represents a non-biased dataset to the best of its ability.

Propensity Score Matching instead takes the propensity scores of each entry and finds the entries with the highest propensity scores within the treatment group and finds the entries within the control group with covariates that most closely match each treatment group entry. This attempts to establish parity between the covariates of the treatment group and the control group.

Both Matching and Stratification work to remove bias from high-dimensional datasets. They do so by balancing out the treated and control groups with processes that emulate a random distribution in an experiment. This is done by evaluating the propensity score of each group. The propensity score represents the probability of the treatment on each sample in the treatment

group and is calculated by mapping the outcomes to a linear regression line. The difference in the methods in how they use the propensity score to balance out the treatment and control group.

In addition, refuters are necessary in the causal analysis process in order to verify the robustness of the results. The following methods all check the effects of confounding variables and compare them with the treatment to concretely establish a causal relationship. Generally, the refuters all involve rerunning the same causal analysis methods with the following changes to the dataset:

- Placebo Refuter: Replaces the treatment variable with a placebo variable with random values
- Data Subset Refuter: Runs the program over a randomly chosen subset of the original data
- Common Cause Refuter: Generates a random confounding variable

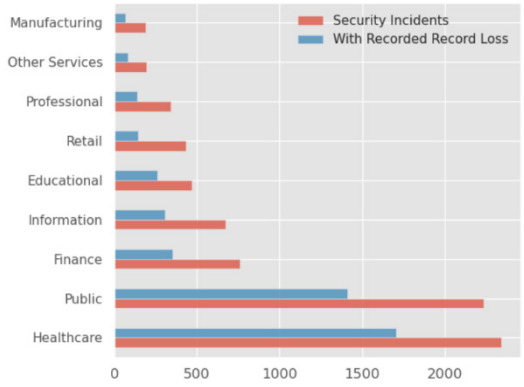


Figure 2: Number of Data Breaches by Industry

3.2. Data Acquisition

We note that high-quality information on real-world cybersecurity incidents through academic or otherwise publicly accessible channels is likely to be unrepresentative of the nature in which breaches occur on a broader scale. As a result, we focus on analyzing healthcare privacy breach data, which generally enjoys stringent reporting standards. Our reasoning is as follows.

For the private sector, disclosure of breaches can negatively impact short term company value as well as consumer trust. A report by IBM's Ponemon Institute in 2019 estimates the global average impact of having a data breach to an organization to be 3.9 million US dollars, representing a rise of 12% over the course of five years. For organizations with fewer than 500 employees, this cost averages to 2.5 million dollars [11]. Voluntary disclosure of data breaches may be unpalatable in light of this [12].

In contrast, government and healthcare institutions are generally under greater legal pressure to disclose similar incidents. For instance, the Department of Human and Health Care Services (HHS) in the United States mandates that information regarding data breaches involving over 500 individuals be disclosed to the media within 60 days of discovery. Structured collection of such breaches is made publicly available through the HHS website [13].

Initial exploration was performed on the VERIS Community Database (VCDB). The VCDB is an open dataset covering a broad spectrum of security incidents occurring throughout both public and private sectors. Data available through this channel represents a small portion of data contained in a more comprehensive report presented in Verizon's annual Data Breach Investigation Report. The VCDB is attractive as there are few publicly available repositories containing annotated security breach information [14].

The VCDB follows the Vocabulary for Event Recording and Incident Sharing (VERIS) framework. Generally, information surrounding security incidents is divided into four categories: Actor, Attribute, Asset, and Action. Actor pertains to the entity or entities responsible for the data breach. Asset characterizes the type of information lost, as well as how accessible said information was. Attribute refers to the degree which the asset in question was affected, as well as the severity of the incident, the medium of transmission, and if said data was exposed to the public. Finally, Action describes how the security breach was carried out; such as if the breach was a result of malware, or simply negligence. Additional data on affected industry and incident timeline are included as well.

To accommodate the wide variety in reporting standards, VERIS uses a fine-grained approach for characterizing security incidents, using a nested key-value store to accommodate some 173 attributes.

We will take the "actor" category as an example. For any given incident, the individual or individuals responsible could be categorized as either external, internal (affiliated with the organization), a partner (associate, but not directly affiliated), or simply unknown or not available. Within each type of actor lies a different subcategory. For instance, the "external actor" label can represent a criminal organization, foreign government, former employee, or a combination thereof. As a result, many of the keys contain lists as values, as represented in (1).

```
"actor": {
        ["external": {
        "variety": ["Mother Nature",
"Criminal Organization"],
                                                    (1)
                "motive": ["NA",
        "Espionage",
                "Ideology"]
        } ...
```

3.3. Why Healthcare?

While the VCDB contains many features describing the companies that were victims of data breach, little information seems to be provided regarding the situation preceding and during the data breach. Therefore, to maintain a degree of uniformity of each company, narrowing down to one industry like healthcare would mitigate discrepancies within the dataset.

Furthermore, the VCDB utilizes a JSON-formatted, hierarchical data structure presented as a list of key-value pairs.

While each record adheres to the same general schema, sparsity arises as a result of how much data is disclosed by each entity, or simply what information is relevant to which sector. Referring back to US Healthcare data breach disclosure law, we can expect a baseline of data to be provided, such as the number of individuals affected, the type of breach, and the vector of attack.

The next logical step was to transform the data from a hierarchical format into a two-dimensional, tabular structure. A strong motivation for this was to make the data both more comprehensible and consistent.

Transforming the database for VCDB was straightforward thanks to the open-source library Verispy. Verispy converts the deeply nested structure of the original VCDB dataset into a two-dimensional grid-like format. by performing "one-hot encoding" on each of the categorical variables.

This leads to a relatively consistent dataset with the caveat of vastly increasing the (perceived) dimensionality. The final table consists of 2,347 columns, containing 2108 (89%) Boolean entries, 147 (6%) string or string-like entries, and the remaining 92 (5%) numerical entries.

1839 entries within VCDB are related to healthcare out of 8363 data breach entries. In order to further narrow down VCDB into healthcare, we further take out all irrelevant variables to our causal model (described in the next section) as well as drop all entries that have empty values in any of those variables. This drops the final dataset entry count to 106 entries, a mere 1.3% of the original VCDB size. This demonstrates sparseness of the VCDB dataset, despite the breadth of information available within.

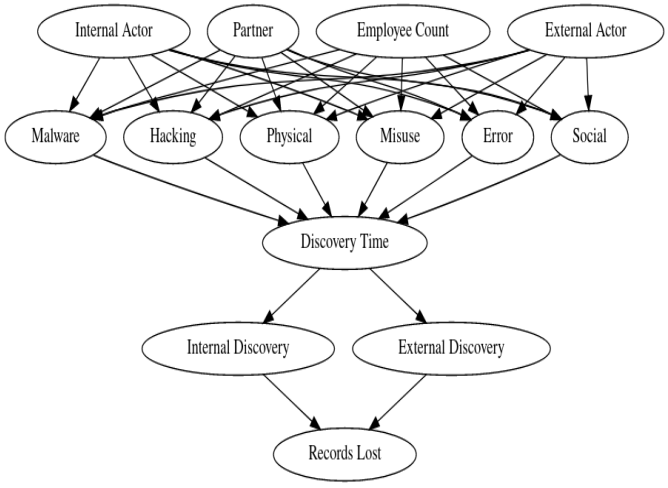


Figure 3. Causal Model for Cybersecurity (on VCDB)

3.4. Data Breach Model

Figure 3 represents the causal graph used as the basis for our causal analysis. The variables are all taken from VCDB and were decided on how accurately they could be mapped to a timeline of the data breach. Since all observational data given to us are all post-data breach, the way to approximate a causal effect for this analysis is to generate a model that shows a progression of events. Many of these variables and their sequencing were derived from personal interpretation than any logical standpoint. We will take a look at each variable type with their justifications.

- Actor
- Employee Count
- Action
- Discovery Time
- Discovery Method
- Records Lost

'Actor' is referring to the one who instigates the action against the victim. This could be a single person, a group of people, or even a natural disaster. In the causal model, the actor is spread amongst three categories: Internal, External, Partner. Internal actors are those who work within the company that is affected by the breach. External actors are those with no affiliation whatsoever with the company. Finally, partners do have or are part of an organization that has an affiliation with the company but are not from the company themselves. This variable represents a general categorizable description regarding the perpetrator of the data breach and is put near the top of the causal graph because the 'actor' is the one that will begin this data breach event.

'Employee Count' represents the general size of the victim company, which is represented by an integer value. Employee count was chosen as it is a variable that conveys a simple, but ordinal description of healthcare organizations.

Each of the types of data breaches ('Malware', 'Hacking', 'Physical', 'Misuse', 'Error', 'Social') are labeled as 'Actions' within VCDB. These are the treatment variables in which the analysis will be performed on. Each action is a binary state, and while there are a few rare cases that have multiple 'Actions' at once, these are still considered one data breach.

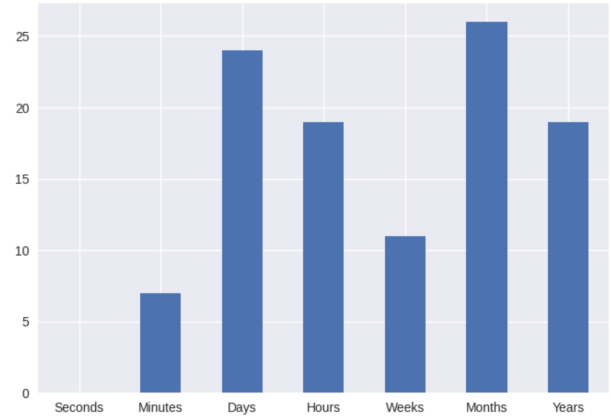


Figure 4. Distribution of 'Discovery Time'

'Discovery Time' is the unit of time it took for the data breach to be discovered. VCDB does not have discovery time as an integer number. Instead, the variable is categorized as six different ranges of numbers, getting subsequently larger. Going under the assumption that the larger unit means that the actual discovery time was longer, the units were combined into one variable from 1-6, each representing a greater scale of time. The unit of time represents the general time frame of the data breach being discovered. Discovery time is one of the outcomes that is used to measure causality of data breaches. 'Containment Time' and 'Exfiltration Time' were considered as well. However, a high

proportion of these entries remain unfilled in the VCDB dataset and are therefore not of much use.

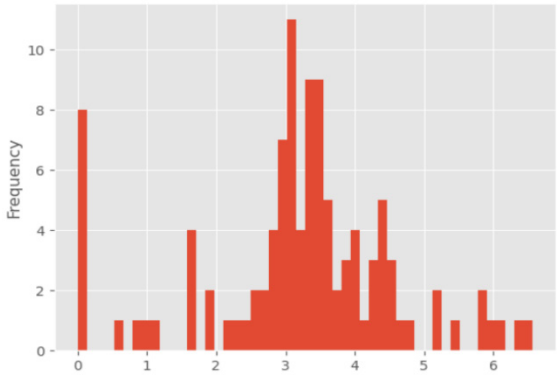


Figure 5. Distribution of Records Lost (log-scaled)

'Discovery Method' is the method by which the victim was first able to discover the data breach. Like 'Actor', this is also split into External, Internal, and Partner, which represents the relationship of the individual or group that discovered the breach to the victim company. External meant those unrelated to the company, Internal part of the company, and Partner are those affiliated with but not directly part of the company.

'Records Lost' is the second outcome we will be using as an outcome to test the causality of the causal model. Similar to Discovery Time, Records Lost is not an integer value, but ranges of values of subsequently greater number. This variable is also similarly combined into one variable ranging from 0-6. One major caveat is that this variable doesn't have a defined unit and thus the scale of a unit of record is determined by each individual company. Part of the decision to focus on healthcare companies only was to mitigate this ambiguity.

4. Results and Analysis

Causal estimate calculations were run across all six 'Actions' (Social, Physical, Misuse, Malware, Hacking, and Error) and two outcomes (Discovery Time, Records Lost). This means multiple runs using the same causal model and dataset but changing the 'Action' and 'Outcome' input for each run until all permutations of each variable was covered. This was then repeated across all refutations. The causal estimate results for Propensity Score Matching on Discovery Time and Records Lost are shown in Figure 6 and 7, respectively.

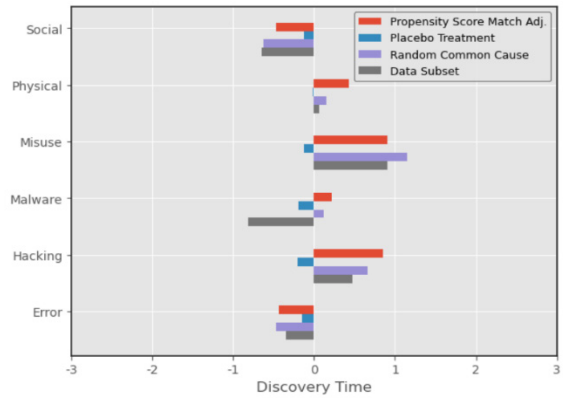


Figure 6. Causal Estimates for Discovery Time

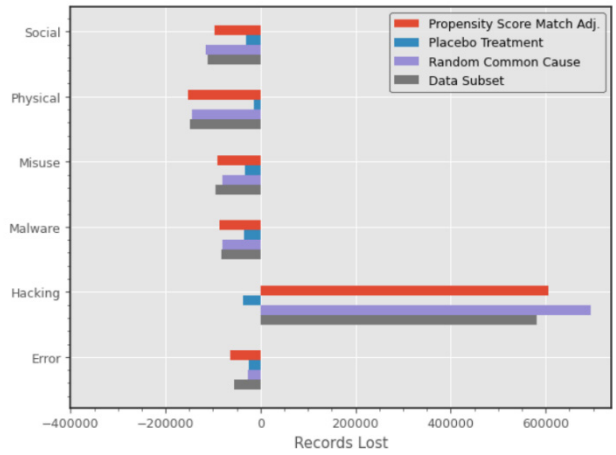


Figure 7. Causal Estimates of Records Lost

As seen in Figure 6 shows, the causal estimate of each action on Discovery Time shows quite a range of values and distributions across all actions. 3 actions (Physical, Misuse, Hacking) have positive causal estimates. indicating a potential strong causal relationship between the actions and lengthy discovery times. On the other hand, Social and Error turn out negative causal estimates, meaning that the impact of those two variables on discovery time is minimal. Lastly, Malware has a unique scenario where there is a split between the Propensity Score and its refuters.

For records lost (Figure 7), Hacking returns an overwhelming higher causal estimate compared to all other actions. In fact, all the other actions return a negative causal estimate. This does not necessarily mean the lack of causal effect of the other actions on records lost. However, it does provide strong indication that the greatest impact when it comes to records lost during a data breach is most likely the result of hacking as opposed to all other methods. Interestingly, this is backed up by both the Random Common Cause and Data Subset, but not the Placebo Refuter. In the Placebo case, the causal estimate returns a comparable negative value to the other actions. A possible explanation can be traced back to the nature of the dataset. While our causal model brings into consideration other causes of data breach, the distribution of the effect of each cause can be hard to separate. This is exacerbated when the Placebo Refuter randomizes the treatment variable, setting it so that every single entry in the dataset can also be considered part of the method of hacking.

This Placebo Refuter discrepancy is reflected across all the actions, which each return strongly diminished causal estimates. However, Hacking remains the only variable where the causal estimate goes from a positive to a negative value.

Another quirk to note is the large value of the causal estimate of Hacking on Records Lost. The reason for this exaggerated value is likely due to a lack of a solid control group within our data. The dataset provides us with a large selection of data breaches in a wide variety of companies. What the dataset lacks are scenarios where no data breach has occurred, generating an inherent bias within the dataset. This bias makes it so that the data do not fit well into linear regression, hence providing an overly large value as the result.

The most unexpected outcome was that propensity score stratification gave inconclusive results when ran on DoWhy, hence the lack of data on this portion of the analysis. After some analysis,

we come to the conclusion that, due to the binary nature of each action, the distribution of the linear regression is not clear enough for stratification to be able to quantify and compartmentalize the dataset into groups. Hence, the resulting value outputs an inconclusive value due to a lack of substantial strata. This applies to stratification and not matching because matching disregards parts of the dataset with low propensity score; in stratification they still have an impact due to those data entries being assigned into strata.

Overall, in this specific scenario and dataset, Hacking would prove to be the most impactful amongst all methods of data breach. However, the refuters give strong indication on where this impact is limited regarding not only the action itself but the dataset as a whole.

5. Conclusion and Future Work

The principal findings of this paper demonstrate the unique perspective of the causal modeling approach. Because we cannot realistically set up an experiment on data breach incidents, particularly in which all factors are readily provided, DoWhy and Causal Modeling allow us to simulate such experiments and make inferences with a degree of robustness based on events that would otherwise be difficult to duplicate.

We identify a subset of factors in the Verizon Community Database and create a hypothesis based on the theory that malware and hacking are the most prominent causes of data breaches. Through propensity score matching and stratification, we measure the strength of the action behind data breaches. By running refutation tests, we are able to verify how well these metrics hold up, similar to how traditional experiments employ control groups or utilize a placebo treatment.

Ample room remains for the use of causal modeling in cybersecurity. We limit the scope of the factors considered in the Verizon Dataset to Actions in order to emphasize the results of the exploratory approach. A larger and denser dataset could utilize the causal model better.

Other fields of cybersecurity lend themselves well to causal modeling. In particular, the use of Directed Acyclic Graphs to model vectors of attack in a network intrusion scenario could lead to different approaches into how such cases are handled.

The study is important to the readers in the scientific community since it is relevant to formulating policies in industry and government, in order to avoid such problems in the near future. Given the context of the work, exhibited in the paper, our findings are worthy of note.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgments

We thank the Data Science Discovery Research Program of UC Berkeley and their participant, Rubina Aujla of UC Berkeley, 2018-2019, for contributions to causal modeling.

References

[1] P. Zornig, Probability Theory and Statistical Applications, De Gruyter. ISBN-13: 978-3110363197
[2] B. McLaughlin, On the Logic of Ordinary Conditionals, Buffalo, NY: SUNY Press, 1990.
[3] J. Pearl, Causality: Models, Reasoning, and Inference, Cambridge: Cambridge University Press, 2000.
[4] J. Pearl, Causality, 2nd edition, Cambridge University Press, 2009.
[5] S. Thornley, R.J. Marshall, S. Wells, R. Jackson, "Using Directed Acyclic Graphs for Investigating Causal Paths for Cardiovascular Disease", Journal of Biometrics & Biostatistics, 2013, 4:182. doi:10.4172/2155-6180.1000182
[6] P. Menzies, H. Beebee. "Counterfactual theories of causation.", 2001.
[7] A. Chesher, A. Rosen, Counterfactual worlds. No. CWP22/15. Centre for Microdata Methods and Practice, Institute for Fiscal Studies, 2015.
[8] A. Agresti, An Introduction to Categorical Data Analysis, 3rd Edition, Wiley Series in Probability and Statistics.
[9] A. Sharma, E. Kıcıman, 2020. Causal Inference and Counterfactual Reasoning. In7th ACM IKDD CoDS and 25th COMAD (CoDS COMAD2020), January 5–7, 2020, Hyderabad, India. ACM, New York, NY, USA,2 pages. https://doi.org/10.1145/3371158.3371231
[10] A. Sharma, E. Kiciman, et al. DoWhy: A Python package for causal inference. 2019.
[11] IBM Security 2019 Cost of a Data Breach Study: Global Overview
[12] R. Anderson, 2001. Why Information Security is Hard-An Economic Perspective. In Proceedings of the 17th Annual Computer Security Applications Conference (ACSAC '01). IEEE Computer Society, USA, 358.
[13] Department of Health and Human Services, 2013. Modifications to the HIPAA Privacy, Security, Enforcement, and Breach Notification Rules Under the HITECH Act and the GINA Act; other Modifications to the HIPAA Rules (78 FR 5565), pp. 5565-5702
[14] VERIZON. Data Breach Investigations Reports Overview, 2019 (DBIR).

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 2, 384-392 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

A Critical Analysis of Usability and Learning Methods on an Augmented Reality Application for Zoology Education

Diego Iquira-Becerra*, Michael Flores-Conislla, Juan Deyby Carlos-Chullo, Briseida Sotelo-Castro, Claudia Payalich-Quispe, Carlo Corrales-Delgado

Universidad Nacional de San Agustn de Arequipa, Perú

ARTICLE INFO

Article history:
Received: 16 January, 2020
Accepted: 11 March, 2020
Online: 07 April, 2020

Keywords:
Augmented reality
Education
Zoology
Usability

ABSTRACT

In recent years, research has been conducted focused on the use of technologies in the classroom, but one of the main problems is to demonstrate that the use of this technology favors learning compared to traditional methods. One of these technologies is augmented reality that allows seeing virtual objects superimposed in the real world, but to achieve its correct use it is necessary to evaluate the usability that is the ease with which an interface is used. In this work we develop and analyze from two perspectives an educational application for zoology that uses augmented reality, a first perspective is usability, where an analysis has been carried out on how the correct design of an augmented reality application should be focusing on heuristic evaluations, the second perspective is at the educational level where we analyze it at the classroom level where we measure the different learning methods (Traditional Learning, Self Learning, and Guided Learning) and which one should be used.

1 Introduction

This paper is an extension of work originally presented in XIII Latin American Conference on Learning Technologies(LACLO) [1]. The emergence of new technologies creates a need to adopt new strategies in the classroom, which complement the learning process to improve the way of teaching.

This research focuses on conducting a case study from the perspective of usability and learning methods of an application of augmented educational reality for the teaching of zoology. Usability is understood as the degree to which users interact with an application that must be effective, efficient and easy to learn [2].

The learning methods have been based on game-based learning where three different methodologies were used (Traditional, Self and Guided) [3].

In the design of the application, we have taken as a basis the traditional method of teaching zoology where cards with descriptions of animals are used, which have been transformed into markers, these markers are recognized by a mobile device that generates a model of the animal and its habitat in 3D. Finally, the case study was carried out in an educational institution where students were divided into groups to which an evaluation was made before and after using the application.

*Diego Alonso Iquira Becerra, Arequipa, Per, diquira@unsa.edu.pe

1.1 Augmented Reality

Augmented reality (AR) is a variation of virtual environments (VE) or virtual reality (VR). The Figure 1 shows a vision of the continuity between virtuality and reality defined by Milgram in 1995 [4], where VE technologies completely immerse the user in a synthetic environment (right of the Figure 1), not being aware of the real world that surrounds it. In contrast, AR allows the user to see virtual objects superimposed in the real world. Therefore, AR complements reality instead of completely replacing it [5].

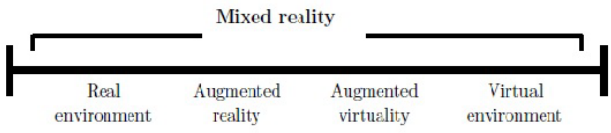


Figure 1: Milgrams reality-virtuality continuum [6]

Azuma defines AR as a technology that meets three characteristics [7]:

1. Combines real and virtual.
2. Interactive in real time.

3. Registered in 3-D.

1.2 Usability

Usability is an attribute that evaluates the ease with which a user interface is used; It can also refer to methods to improve ease of use during the design process [2].

The 5 main components of usability are defined by the following questions:

1. Learning: How easy is it for users to perform basic tasks the first time they encounter the design?
2. Efficiency: Once users have learned the design, how quickly can they perform tasks?
3. Easy to remember: When users return to the design after a period of not using it, how easily can they restore their competition?
4. Errors: How many errors do users make, how serious are these errors and how easy can they recover from errors?
5. Satisfaction: How nice is it to use the design?

. An important point is utility, which refers to the functionality of the design: Does it do what users need? [2].

ISO 9241 is a standard focused on quality, providing requirements and recommendations relating to the attributes of the hardware, software, and environment that contribute to the usability and ergonomic principles that guide them [8].

Specifically, ISO 9241-11 tells us that the objective of the design and evaluation of systems, products, and services for usability is to allow users to reach goals effectively, efficiently and with satisfaction, taking into account the context of use. Thus, it defines usability as "The extent to which a system, product or service can be used by specific users to achieve specific objectives with effectiveness, efficiency, and satisfaction in a specific context of use" [8] [9].

Usually, to measure usability according to ISO 9241-11, it is done by user tests, the advantage of these is that resources can be quite modest. The number of users varies, but a small sample of 5 to 8 users can identify 80-85% of usability problems in a controlled environment.

The goal of usability evaluation is to see the degree to which a system is easy and pleasant to use, to determine the scope of its usability, using robust, reliable and objective metrics [10]. The evaluation of traditional systems is based on predefined usability methods. These can be classified as Inspection methods, test methods and user reports [10]- [12].

1.3 Learning Methods

1.3.1 Traditional Learning

In the traditional teaching method participants are required to attend the classes in the fix hour and place regularly and receive training face to face, mainly in the form of verbal and textual ones. What is more remarkable is that this method of teaching is provided in the same way for all learners [13].

Therefore, this method generates space-time limitations since it is only performed at predetermined times and is limited to only being carried out in classrooms [14].

1.3.2 Self Learning

According to Acosta, self learning is a form of active learning in which students, guided by a facilitator, autonomously increase their skills in a specific area [15].

Operational units of students or autonomous work teams are established, which choose a situation or reality to investigate by team, then establish their baseline of competencies, then choose their learning objectives, finally design a simulation scenario, experience it and reflect. Likewise, self-learning can be understood as the process in which the student is involved in the construction of their learning, thus promoting autonomous work [16].

1.3.3 Guided Learning

Guided learning lies midway between direct instruction and free learning, presenting a learning goal, and scaffolding the environment while allowing students to maintain a large degree of control over their learning [17].

Guided learning approaches are effective because they create learning situations that encourage students to become active and engaged partners in the learning process [17].

1.4 Zoology Education

Children from birth have a scientific interest in exploring and knowing the objects and beings of their environment, for this reason, the teaching of science and technology encourages their curiosity, discovery and the pleasure to learn, as well as providing sensitivity and respect for the environment and living things, specifically in the area of zoology which focuses on animal studies.

In the first years of age, children develop their ability to understand nature in their environment, where the student questions, reflects and thinks about the beings that inhabit it, for this reason, the Ministry of Education of Peru (MINEDU) proposes the use of learning paths to guide the correct teaching of students [18].

So that the student can progress in the learning pathways of the zoology area, it is expected that they will be able to explain the beings of nature-based on evidence from various sources of information and that they will be able to describe in their own words the characteristics or qualities of living beings in their environment based on previous knowledge.

According to MINEDU, the learning paths for zoology teaching are the competences, capacities, and indicators which focus on [18]:

- Competence: It is intended that the student explain the physical world based on their scientific knowledge and that their scientists understand and apply knowledge, arguing them scientifically.
- Capacity: The student must be able to recognize animals according to their characteristics and must be able to relate an animal according to its environment.

- Indicators: The student must describe the characteristics of the animals, must mention their similarities between their descendants and parent and must relate them according to their environment.

Once the learning path is completed; students will be able to: learn the characteristics and needs of living beings, understand that living beings are born from another living being, grow and develop, similarities between parents and descendants and the names of Animals and the place they inhabit.

Another central point in the learning paths is the map of progress that are instruments that allow identifying the progress of the students, to have greater clarity regarding the expectations that are expected of them regarding each established competence [18].

At a more specific level, the progress maps contain a clear and agreed definition of the learning goals that must be achieved by all students at the end of a learning cycle, in this way we can determine that the progress maps provide us with all the information necessary of the students according to each competition [19].

Finally, the materials used in the classes are [18]:

- Metal board and magnetized animal parts.
- Cards with descriptions of the animals.
- Paper, markers and duct tape.
- Science and Environment Book provided by MINEDU.

2 Related Works

In recent years, interest in the use of augmented reality for teaching in children has increased. Different previous investigations have built applications through programs such as Unity (Multi-platform video game engine) used in conjunction with augmented reality libraries such as Vuforia. A study [20] shows a mobile application of augmented reality to reinforce student learning and experience, among the activities of the application was the 3D visualization of the Earth by focusing the view on the encyclopedia and being able to visualize an animal model in 3D using as markers bills from different countries. All this allowed increasing the curiosity of the student by stimulating his memory since he related the virtual 3D models with the physical elements.

In another study, they developed educational software in Unity for teaching physics where a methodology for the evaluation of didactic usability was proposed [21].

Another study [22] is about how to create educational software using augmented reality combined with Bloom's Taxonomy, whose main purpose was to stimulate spatial reasoning by improving the children's learning process. In another investigation [23], an application has been created which takes a specific image of an anatomy atlas and converts it into a three-dimensional and dynamic model , which allows improving the interactivity of the student in the learning process, in that system they are stored in a database the different scanned photos of the anatomy atlas and then are consulted by the application, they made the three-dimensional models in Unity3D and Maya, one of the most important points of that research is that it encourages students' self-learning.

Teaching an alphabet can be complicated at times, but augmented reality also works for these cases, as in this study [24], where an attractive and interesting learning process of the Hijaiyah alphabet is developed through a mobile application using augmented reality technology. The purpose of this application is to encourage children's interest in learning the Hijaiyah alphabet. This application uses smartphones and markers. It was built using Unity and the Vuforia library, as well as Blender for modeling 3D objects. To use the application, the smartphone camera tracks the marker and once it is identified, the marker will have projected the objects of the Hijaiyah alphabet in three-dimensional form. Finally, the user can learn and understand the form and pronunciation of the Hijaiyah alphabet by touching the virtual button on the marker.

Augmented reality can also be used in the area of chemistry as in the application called PeriodikAR [25], which is an application developed for the Android platform through the use of the Vuforia library. The tracking technique used is without a marker. This application runs dynamically to provide certain information about the elements in an animated video. The information includes the name of each element, atomic number, boiling point, melting point, density, atomic mass, standard atomic weight, oxidation state, symbol, phase, element category, electron configuration, and electron orbital display. In this case, the dynamic term means that all the markers and animated videos of the application can be added or subtracted without modifying the source code of the application. It is because all the markers and the storage of animated videos are placed in a different place from the application storage. This application aims to increase interest in studying chemical elements. Finally, another work [26] focuses on an Android application that teaches playing musical instruments. This application was made using technologies such as augmented reality Vuforia, 3D modeling with Blender and Unity 3D.

3 Zoology Augmented Reality Application

3.1 Application Overview

The application was developed for mobile devices that have a camera, which recognizes the markers found on the cards; the markers contain images of the animals in their environment. The target audience is kindergarten students, who can interact with the virtual animals that appear in their habitat.

The objective of the application is to provide new technological tools to the classroom, specifically the cards with virtual reality markers, which will allow students to interact and recognize the animals according to their characteristics and their environment. Additionally, activities in the application focused on self-learning were added, where the student recognizes the animals based on their shadow, these activities contain a system of scores and notifications; in order to generate continuous feedback between the student and the application.

3.2 Hardware and Software

For the development of the application, we have used Android mobile devices version 7.0 or higher, which had a gyroscope. Also, a computer with the following features was used: Intel Core I7

processor, 8 GB of RAM, 1 TB of storage memory and an Nvidia GeForce GTX 980M graphics card.

The programs that were used were: Unity 3D and Vuforia for the main development of the application, Inkscape for graphic design and Audacity for audio editing.

3.3 Activities

The application will teach about the animals of the following groups: mammals, birds, reptiles, fish and amphibians as shown in Figure 2, the application will have five levels that represent the learning sessions of each day: Monday, Tuesday, Wednesday, Thursday and Friday, as seen in Figure 3, we will describe the activities of each level below:

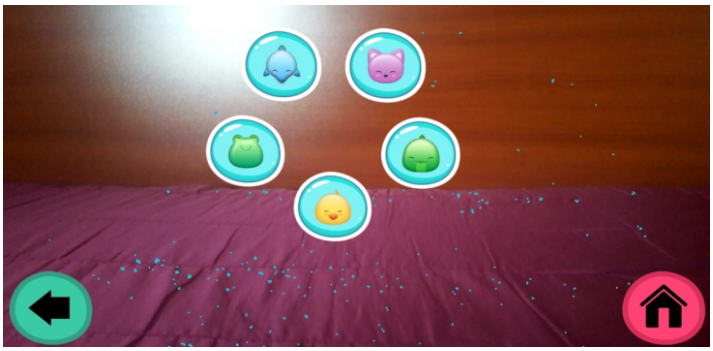


Figure 2: Groups of animals: mammals, birds, reptiles, fish and amphibians (Source: Own Elaboration)

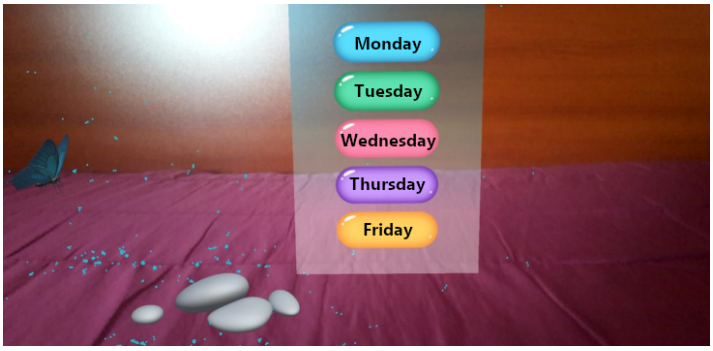


Figure 3: Application levels: Monday, Tuesday, Wednesday, Thursday and Friday (Source: Own Elaboration)

3.3.1 Monday

It is the first level of the lessons, in which students learn the names and descriptions of each of the animals. They can also observe the models of each of the animals.

3.3.2 Tuesday

In this level, the animations of each animal are presented, which are interactive through the use of touch.

3.3.3 Wednesday

At this level, 3d animal models are presented in a 3d environment that represents their habitat as seen in Figure 4.

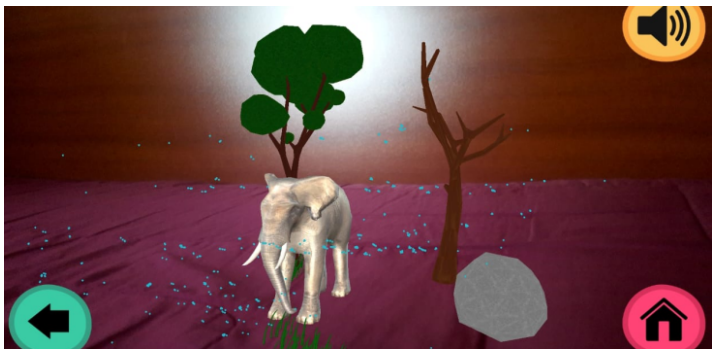


Figure 4: 3d animal models in their habitat (Source: Own Elaboration)

3.3.4 Thursday

This level shows the 3d animal models that were seen in the previous levels but on a real scale, as seen in Figure 5.

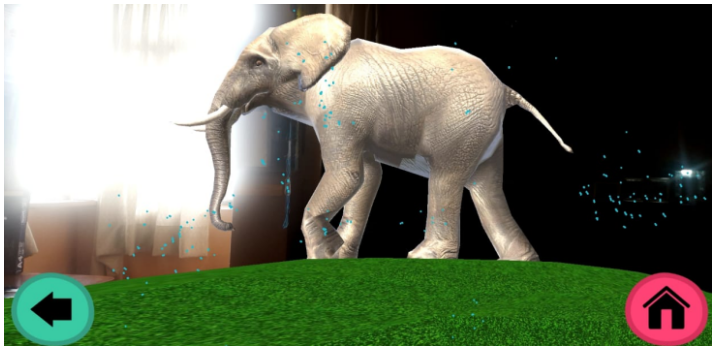


Figure 5: Real-size animal 3d model with animations (Source: Own Elaboration)

3.3.5 Friday

Final level, the student link the animals to their habitat, which are selected randomly. The student must drag the animal with its corresponding habitat, see Figure 6.



Figure 6: Activity to relate animals with their habitats (Source: Own Elaboration)

4 Methodology

4.1 Usability Evaluation

To perform the evaluation, we made use of user tests that contain a combination of task completion rates, errors, task times and task level satisfaction. Two ways of testing were considered:

- Formative: Provides immediate feedback to improve learning
- Summative: Evaluate what has been learned.

The experiences to evaluate and improve user interaction with the application was formative, and the problems were quantified in terms of frequency and severity, we monitored the users who had problems, we measured the time it took to complete a task and we determined whether they completed the tasks correctly. The method used to evaluate the results is described below.

1. Task Time: The time a user spends in an activity, specifically:
   - Completion task time: Time for users to complete the task correctly.
   - Failure task time : Time for users to complete the task incorrectly.
   - Total task time: Total duration of time users spend on a task.
2. Errors: The capture of touch events in the activities of the application, considering actions such as:
   - ACTION DOWN: This event is manifested when the screen is pressed and we don't move the pointer anywhere.
   - ACTION MOVE: This event is manifested when after pressing the screen we proceed to move the pointer, in other words, when we do these 2 events will appear, the ACTION DOWN and then when moving the pointer the ACTION MOVE event will appear.
   - ACTION UP: This event is manifested when we lift the finger or pointer of the screen, for this event to occur, the ACTION DOWN event must have occurred before.
   - ACTION OUTSIDE: This event is manifested when the event occurs outside the normal limits of an user interface element.
3. Effectiveness: Measured by the rate of completed tasks and the rate of errors in tasks, the following will be analyzed: Completed tasks, Objectives achieved, Errors in a task, Tasks with errors and Intensity of task error.
4. Efficiency: It is measured by the time it takes to complete a task, it will be analyzed: Task time, Time efficiency, Profitability, Production time ratio, and Unnecessary actions.
5. Satisfaction: The perception of the ease of use of a system is measured, if they can be completed immediately after a task (post-task questionnaires), at the end of a usability session (post-test questionnaires), or outside of A usability test is analyzed.

4.2 Evaluation Learning Methods

4.2.1 Participants

A total of 45 5-year-old children participated in the study, they were kindergarten students of a national school; There were a greater number of men (62.2 %) than women, the students belonged to two classrooms.

4.2.2 Method

The students were divided into three groups chosen at random, they were given the first test to measure their knowledge about zoology, this test focused on presenting a group of thirteen animal cards, of which the students had to recognize the name and habitat of different animals.

Subsequently, each group was assigned a different teaching method, a first group used the traditional learning where the teacher used cards with descriptions of the animals to teach them about zoology, the second group were given mobile devices and a set of markers that contained figures of animals with which students should self-learning about animals, finally, the third group used mobile devices and markers but their learning was led by a teacher who guided and explained about the different animals.

After the end of the learning session, a second evaluation was carried out on the students in which the acquired knowledge was measured.

5 Results

5.1 Usability Results

To evaluate the results, three aspects of usability were considered: Effectiveness, efficiency, and satisfaction. We capture the events of user interaction with the application in Friday activity.

The measurement by task completion rate and task error rate are shown in Figure 7.

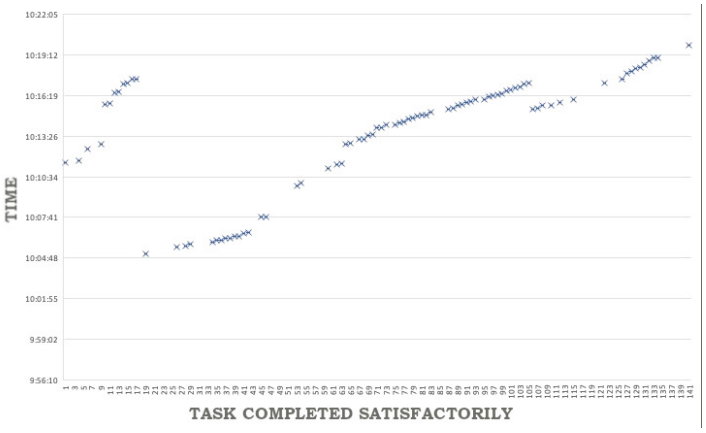


Figure 7: Tasks completed satisfactorily (Source: Own Elaboration)

To analyze the results, the tasks completed in the sample of 141 tasks were observed as a collection of points in time, the position of each point on the vertical axis (y) is the time at which the value

was collected. Tasks that do not have a point associated were not completed successfully.

To recognize effectiveness as a characteristic of the usability evaluation of the application, the total percentage of tasks completed with or without cognitive errors related to animals was considered. Figure 8 shows that 91% represents the total of complete tasks with or without cognitive errors, while 9% represents tasks that did not reach the objective. Comparing the result with the percentage of tasks completed, 31% represents tasks completed with cognitive errors.

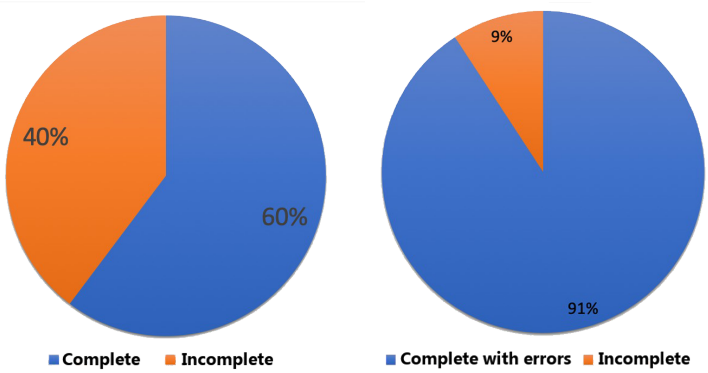


Figure 8: Percentage of completed tasks and complete tasks with errors (Source: Own Elaboration)

To recognize the efficiency, we considered the task time in the accomplishment of the task and an average time of 5 seconds was considered. Of the sample size of 141 tasks, only the time of 128 tasks that were completed was considered. In Figure 9, the time of accomplishment of each task is observed; The average time ranges from 0 seconds to 5 seconds, considering the highest peak with 23 seconds and the lowest peak 0 seconds.

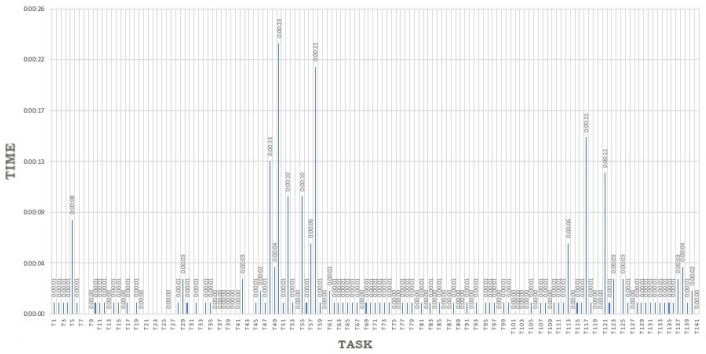


Figure 9: Task Time (Source: Own Elaboration)

In Figure 10, it is observed that the tasks completed within the average time of 5 seconds represent 91%, being a total of 116 tasks. While the remaining 9% represents a total of 12 tasks that obtained a task completion time greater than 5 seconds. According to the evaluation of this characteristic of Usability, the high percentage of tasks completed within the average time assigned is observed, therefore, it can be affirmed that the application in Friday activity complies with this aspect of usability by 91%.



Figure 10: Percentage of the execution time of a task (Source: Own Elaboration)

To know satisfaction, we measured with the SUS (System Usability Scale) questionnaire, an instrument of Likert scale of 10 items. The evaluation was carried out immediately after the interaction, which allowed the evaluators to record their initial feelings and responses with greater certainty.

According to the results presented in Figure 11, the SUS score obtained is 69. Since the theoretical maximum is 100, we can deduce good but not excellent usability of the system.

SUS QUESTIONNAIRE												
RATING FROM 1 TO 5:												
1: TOTAL DISAGREE												
5: TOTAL AGREE												
	QUESTIONS	E1	E2	E3	T1	T2	PSY	Al	A2	A3	A4	Pts
1	I think I would use this application frequently	5	5	5	5	4	4	4	4	4	5	4.5
2	I find this application unnecessarily complex	1	1	1	1	1	1	1	1	1	1	1
3	I think the application was easy to use	4	4	4	4	5	5	5	5	5	4	4.5
4	I think I need help from a person with technical knowledge to use this application.	3	3	3	3	4	4	4	4	4	3	3.5
5	The functions of this application are well integrated	4	4	4	4	4	4	4	4	4	4	4
6	I think the application is very inconsistent	2	2	2	2	1	1	1	1	1	2	1.5
7	Most people would learn to use this application very quickly.	5	5	5	5	5	5	5	5	5	5	5
8	I find that the application is very difficult to use	1	1	1	1	2	2	2	2	2	1	1.5
9	I feel confident using this application.	5	5	5	5	4	4	4	4	4	5	4.5
10	I needed to learn many things before being able to use this application.	2	2	2	2	2	2	2	2	2	2	2
E= Engineer T= Teacher PSY= Psychologist A= Education Assistant												

Figure 11: SUS questions of the application (Source: Own Elaboration)

5.2 *Results Learning Methods*

To analyze the results obtained from the pre-test and the post-test we decided to apply the regression and correlation of data, which were applied to both the names of the animals and the habitat of the animals of the following thirteen animals: cow, shark, frog, penguin, polar bear, lion, giraffe, gorilla, seal, elephant, crocodile, sea horse, and owl.

In the Figure 12, the data was processed with the STATA (Software for Statistics and Data Science) tool where a regression equation was used that explains the students' learning about the names of the animals which turned out to be in inverse function of the self-learning method, that is that by applying the method of

self-learning, the learning of animal names is reduced but there is a direct relationship between guided learning and traditional learning.

A regression equation was used for the analysis of animal name learning, this equation is a function of self-learning plus guided learning towards traditional learning.

```
. regress ANIMAL NAME  SELF LEARNING  GUIDED LEARNING  TRADITIONAL LEARNING

      Source |       SS       df       MS              Number of obs =      45
-------------+------------------------------           F(3, 41)      =   41.17
       Model |  6.94039066     3  2.31346355           Prob > F      =  0.0000
    Residual |  2.30405378    41  .056196434           R-squared     =  0.7508
-------------+------------------------------           Adj R-squared =  0.7325
       Total |  9.24444444    44  .21010101           Root MSE      =  .23706

------------------------------------------------------------------------------
ANIMAL NAME |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------+-----------------------------------------------------------------
 SELF LEARNING |  -.0088418   .0183371    -0.48   0.632    -.0458744    .0281908
 GUIDED LEARNING |   .1384828   .0125899    11.00   0.000     .1130569    .1639087
TRADITIONAL LEARNING |   .0203049   .0201334     1.01   0.319    -.0203553    .0609652
       _cons |  -.5299198   .1554088    -3.41   0.001    -.8437742   -.2160653
------------------------------------------------------------------------------
```

Figure 12: Regression of animal name learning between different learning methods (Source: Own Elaboration)

$$LEARNING\quad ANIMAL\quad NAME^{(\wedge=)}$$
$$-0.0088418\quad Self\ learning$$
$$+0.1384828\quad Guided\ learning$$
$$+0.203049\quad Traditional\ learning + u$$

The negative self-learning coefficient indicates that it has an inverse relationship with the learning of animal names, that is, when the learning of animals is achieved, self-learning has little influence. On the other hand, guided learning and traditional learning have a positive sign in their coefficients, which indicates that they influence the learning process.

For this reason the model is significant in its entirety, since its p-value of F is 0.0000 which is lower than the significance level of 5%.

The independent variables (Self learning; Guided learning; Traditional learning) explain the variable learning of animal names by 75.08%.

With which we determine the following:

- The self learning variable is not significant for the model since the p-value of its coefficient is 0.632 which is less than 5%.
- The guided learning variable is significant for the model since the p-value of its coefficient is 0.000 which is less than 5%.
- The traditional learning variable is not significant for the model since the p-value of its coefficient is 0.319 which is greater than 5%

In the same way, we perform another analysis with the data corresponding to the names of the animals where it is stated that the variable that has the greatest correlation with the learning of the name of the animals as seen in the Figure 13 is the guided learning with 86.21 % and the variable that has a negative correlation of -5.99% is the traditional learning.

```
. corr ANIMAL NAME  SELF LEARNING  GUIDED LEARNING  TRADITIONAL LEARNING
(obs=45)

                |     NAME     SELF   GUIDED  TRADITIONAL
----------------+------------------------------------------
     ANIMAL NAME |   1.0000
   SELF LEARNING |  -0.0599   1.0000
  GUIDED LEARNING |   0.8621  -0.0263   1.0000
TRADITIONAL LEARNING |   0.1079   0.0031   0.0342   1.0000
```

Figure 13: Correlation of animal name learning between different learning methods (Source: Own Elaboration)

On the other hand, we perform another regression equation was used for the analysis of animal habitat learning, this equation is a function of self-learning plus guided learning towards traditional learning. 14.

```
. regress ANIMAL HABITAT  SELF LEARNING  GUIDED LEARNING  TRADITIONAL LEARNING

      Source |       SS       df       MS              Number of obs =      45
-------------+------------------------------           F(3, 41)      =   49.14
       Model |  8.27600752     3  2.75866917           Prob > F      =  0.0000
    Residual |  2.30177026    41  .056140738           R-squared     =  0.7824
-------------+------------------------------           Adj R-squared =  0.7665
       Total |  10.5777778    44  .24040404           Root MSE      =  .23694

------------------------------------------------------------------------------
ANIMAL HABITAT |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------+-----------------------------------------------------------------
 SELF LEARNING |   .0153236   .0201697     0.76   0.452    -.0254098    .0560571
 GUIDED LEARNING |   .1841973   .015174    12.14   0.000     .1535529    .2148417
TRADITIONAL LEARNING |  -.0155921   .0183617    -0.85   0.401    -.0526744    .0214902
       _cons |  -.7618375   .1579508    -4.82   0.000    -1.080826   -.4428494
------------------------------------------------------------------------------
```

Figure 14: Regression of animal habitat learning between different learning methods (Source: Own Elaboration)

$$LEARNING\quad ANIMAL\quad HABITAT^{(\wedge=)}$$
$$0.0153236\quad Self\ learning$$
$$+0.1841973\quad Guided\ learning$$
$$-0.0155931\quad Traditional\ learning + u \tag{1}$$

The negative traditional learning coefficient indicates that it has an inverse relationship with the learning of animal habitat, that is, when the learning of animals habitat is achieved, traditional learning has little influence. On the other hand, guided learning and self learning have a positive sign in their coefficients, which indicates that they influence the learning process.

The model is significant in its entirety since its p-value of F is 0.0000 which is below the significance level of 5%. The independent variables (Self learning; Guided learning; Traditional learning) explain the variable learning of the location of the animals by 78.24%.

With which we determine the following:

- The self learning variable is not significant for the model since the p-value of its coefficient is 0.452 which is less than 5%.

- The guided learning variable is significant for the model since the p-value of its coefficient is 0.000 which is less than 5%.
- The traditional learning variable is not significant for the model since the p-value of its coefficient is 0.401 which is greater than 5%

Additionally, we perform another analysis with the data corresponding to the habitat of the animals as seen in Figure 15, where it is stated that the variable that has the greatest correlation with the learning of the location of the animals is the guided learning with an 88,07% and the variables that have a negative correlation of -1.62% and -0.74% are those of self and traditional learning respectively.

```
. corr ANIMAL HABITAT SELF LEARNING GUIDED LEARNING TRADITIONAL LEARNING
(obs=45)
```

	HABITAT	SELF	GUIDED	TRADITIONAL
ANIMAL HÁBITAT	1.0000			
SELF LEARNING	-0.0162	1.0000		
GUIDED LEARNING	0.8807	-0.0792	1.0000	
TRADITIONAL LEARNING	-0.0074	0.0229	0.0600	1.0000

Figure 15: Correlation of animal habitat learning between different learning methods (Source: Own Elaboration)

Finally, we made a comparison between the pre-test and post-test of the different learning methods as seen in Figure 16, in which the guided learning was the one that obtained the highest percentage of improvement, they get to improve the understanding of animals and the habitat in 48.6% followed by self learning with a 30.5% improvement and finally traditional learning that obtained only 21%.

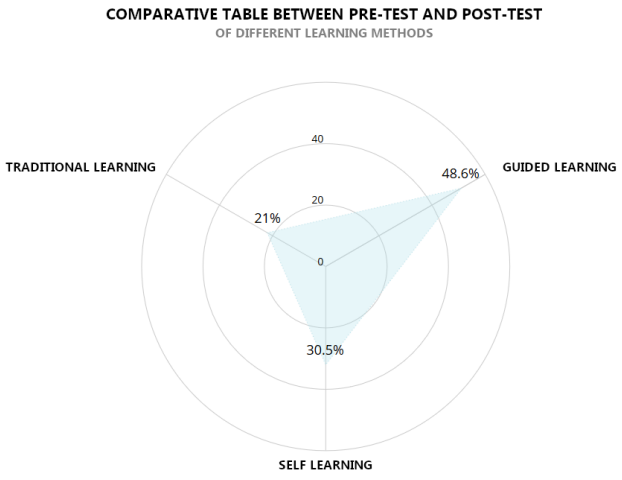


Figure 16: Comparative table between pre-test and post-test of different learning methods (Source: Own Elaboration)

6 Conclusions and Future Works

6.1 Conclusions

According to the usability characteristics established by ISO 9241-11 presented in this research, it provides a practical way to evaluate the usability of AR applications. Thus, the proposal is expected to become part of the AR application development process, which will allow users to perform tasks that achieve the objectives of effectiveness, efficiency, and satisfaction.

During the evaluation carried out in the development of the applications, several design problems were found including the use of AR techniques in the application. When the end-users of the application are children, it is advisable not to use markers for the recognition of 3D elements, it was difficult to keep their attention on pointing to the marker and at the same time seeing the recognition image. Due to these first tests, the application uses navigation and selection techniques. Therefore, it is concluded that to choose the interaction techniques depends on the requirements of the application. The same set of techniques will not work well in all situations.

An important observation was that children seemed to understand how to use the augmented reality application as soon as they began to interact with technology.

In the process of development and evaluation of the application, it is essential to involve teachers in the development of activities directly related to the needs of students, especially if the application is aimed at children 5 years old; since it allows researchers to have a better overview and better understand the problems that could occur.

We can conclude that the applications of AR that follow the principles of usability help in the motivation and learning of children.

When analyzing the results obtained in the tests we determine that for the recognition of animal names a guided learning method turns out to be better than a self-learning.

Which was because the application failed to explain the name correctly, since it was not repeated and when it was heard at the same time that the sound generated by the animal caused a distraction in the students and did not take the corresponding attention unlike of guided learning, where the teacher indicated the animal they should look for and explained information about the animal that is complemented by the visualization of the 3D model.

As for the habitat, the guided learning method turned out to be the best, because with traditional learning the student could not identify visually the environment where the animal lived, since the habitat was described orally, on the other hand in the car if the student could recognize the environment, but since there were no dialogues that would indicate information about the environment, the student had to rely on his previous knowledge to identify the correct name of the environment, finally when comparing the post-test and the pre-test we identified that the guided learning method obtained a higher score reaching 48.3% of positive responses.

With which we have been able to conclude that the use of augmented reality does show improvements in traditional education, but the best method to carry out education through this technology is that the teacher guides the student on what to do and can maintain at all times the control of the actions that each student will take.

6.2 Futures Work

Articles focused on the use of augmented reality in the classroom will continue, while many will focus on short-term research, it

is also necessary to consider long-term research and the use of other technologies, so we suggest the following topics that could be investigated:

- Add specific ergonomic tests regarding the user experience. It is also recommended to consider ergonomics usability issues in terms of interaction with mobile devices.
- Perform more extensive tests using the evaluation proposal presented to evaluate new virtual contents and the use of other types of 3D interactivity techniques or sensors.
- Perform new educational activities that focus on Machine Learning to improve the free education process and serve as a complement to targeted education.
- Find new technologies of augmented reality that not only focus on markers such as ARCore for the creation of new educational activities focused on the topic of zoology.
- The creation of other augmented reality applications that explain natural phenomena that occur in the student environment is proposed.

Acknowledgment Thanks to the Universidad Nacional de San Agustn de Arequipa for support in the research.

References

[1] D. A. I. Becerra, B. D. S. Castro, M. M. F. Conislla, and C. Corrales-Delgado, "Augmented reality applied in the design of learning activities in zoology," in *2018 XIII Latin American Conference on Learning Technologies (LACLO)*. IEEE, 2018, pp. 121–126.

[2] J. Nielsen, "The usability engineering life cycle," *Computer*, vol. 25, no. 3, pp. 12–22, 1992.

[3] E. Danniels and A. Pyle, "Definir el aprendizaje basado en el juego," *Aprendizaje basado en el juego. Actualizado en febrero de*, 2018.

[4] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, "Augmented reality: A class of displays on the reality-virtuality continuum," in *Telemanipulator and telepresence technologies*, vol. 2351. International Society for Optics and Photonics, 1995, pp. 282–292.

[5] S. E. Schaeffer, "Usability evaluation for augmented reality," 2014.

[6] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE TRANSACTIONS on Information and Systems*, vol. 77, no. 12, pp. 1321–1329, 1994.

[7] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoperators & Virtual Environments*, vol. 6, no. 4, pp. 355–385, 1997.

[8] M. Kurosu, "Nigel bevan and concepts of usability, ux, and satisfaction," *Journal of Usability Studies*, vol. 14, no. 3, 2019.

[9] M. Georgsson, N. Staggers, E. Årsand, and A. Kushniruk, "Employing a user-centered cognitive walkthrough to evaluate a mhealth diabetes self-management application: A case study and beginning method validation," *Journal of biomedical informatics*, vol. 91, p. 103110, 2019.

[10] G. Cockton, "Usability evaluation," 2012.

[11] H. Martínez and P. Bandyopadhyay, "Analysis of four usability evaluation methods applied to augmented reality applications," *Helsinki, Finland: Helsinki Institute for Information Technology (HIIT)*, 2014.

[12] O. Sharhorodska and D. Iquira, "Interaction of low cost mobile virtual reality environments–using metaphor in an astronomy laboratory," in *International Conference on Human-Computer Interaction*. Springer, 2019, pp. 293–300.

[13] M. Gharibpoor, S. Sargazi, and M. Aref, "Efficiency evaluation of e-learning compared to traditional education in human resource development (case study: Small and medium enterprises in shiraz)," in *7th International Conference on e-Commerce in Developing Countries: with focus on e-Security*. IEEE, 2013, pp. 1–8.

[14] J. E. Thiele, "Learning patterns of online students," *Journal of Nursing Education*, vol. 42, no. 8, pp. 364–366, 2003.

[15] C. L. Costa, A. M. Nicolás, J. A. G. Méndez, M. d. G. A. Martínez, and J. L. D. Agea, "Enseñando con metodología de autoaprendizaje en entornos simulados (maes©). un estudio cualitativo entre profesores y alumnos de grado en enfermería," *Educación Médica*, vol. 20, pp. 52–58, 2019.

[16] C. J. Z. Rodríguez, P. García-Martínez, C. Ferreira, I. Fernández, and J. J. Miret, "Sistemas interactivos eficientes para la tutorización y el autoaprendizaje de óptica basados en la resolución gráfica de problemas," *Optica pura y aplicada*, vol. 46, no. 2, pp. 197–203, 2013.

[17] D. S. Weisberg, K. Hirsh-Pasek, and R. M. Golinkoff, "Guided play: Where curricular goals meet a playful pedagogy," *Mind, Brain, and Education*, vol. 7, no. 2, pp. 104–112, 2013.

[18] MINEDU. (2016) Programa curricular educacin inicial. [Online]. Available: http://www.minedu.gob.pe/curriculo/pdf/programa-curricular-educacion-inicial.pdf

[19] M. J. Ramos Haro, G. García Figueroa, E. Y. Quinteros Hijar, M. J. Niño Correa, M. Ventura Panduro, W. Palomino Noa, J. M. Camavilca Vega, J. A. Pezo de la Cuba, L. Velasco Taipe, C. Yupán Cárdenas *et al.*, "Rutas del aprendizaje¿ qué y cómo aprenden nuestros niños y niñas?: Ii ciclo. área curricular ciencia y ambiente, 3, 4 y 5 años de edad. versión 2015," 2015.

[20] "An educational augmented reality app to facilitate learning experience." *2017 International Conference on Computer and Applications (ICCA), Computer and Applications (ICCA), 2017 International Conference on*, p. 279, 2017.

[21] D. Alonso Iquira Becerra, J. Herrera Quispe, R. Guillermo Apaza Aceituno, G. Mary Poma Vargas, F. Gabriela Fernandez Zamora, J. Luis Huillca Mango, G. Paulina Anccasi Figueroa, A. Alexis Perez Vizcarra, and J. Willian Torres Chana, "Evaluation of a gamified 3d virtual reality system to enhance the understanding of movement in physics," 01 2017, pp. 395–401.

[22] "Playful and interactive environment-based augmented reality to stimulate learning of children." *2016 18th Mediterranean Electrotechnical Conference (MELECON), Electrotechnical Conference (MELECON), 2016 18th Mediterranean*, p. 1, 2016.

[23] J. A. Juanes, D. Hernández, P. Ruisoto, E. García, G. Villarrubia, and A. Prats, "Augmented reality techniques, using mobile devices, for learning human anatomy," in *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*, ser. TEEM '14. New York, NY, USA: ACM, 2014, pp. 7–11.

[24] R. F. Rahmat, F. Akbar, M. F. Syahputra, M. A. Budiman, and A. Hizriadi, "An interactive augmented reality implementation of hijaiyah alphabet for children education," in *Journal of Physics: Conference Series*, vol. 978, 2018.

[25] A. A. K. Oka Sudana, A. Setiawan, and I. P. A. E. Pratama, "Augmented reality for chemical elements: Periodikar," *Journal of Theoretical and Applied Information Technology*, vol. 90, no. 1, pp. 88–92, 2016, cited By :2.

[26] M. K. Zhaparov and U. Assanov, "Augmented reality based on kazakh instrument "dombyra"," in *8th IEEE International Conference on Application of Information and Communication Technologies, AICT 2014 - Conference Proceedings*, 2014.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 6, 386-396 (2019)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Novel Cost Function based Motion-planning Method for Robotic Manipulators

Dániel Szabó*, Emese Gincsainé Szádeczky-Kardoss

Department of Control Engineering and Information Technology, Budapest University of Technology and Economics, H-1117, Hungary

ARTICLE INFO

Article history:
Received: 09 October, 2019
Accepted: 01 December, 2019
Online: 30 December, 2019

Keywords:
robotic manipulator
motion-planning
Rapidly-exploring Random Tree
Fuzzy-approximation

ABSTRACT

In this paper an offline motion-planning algorithm is presented for robotic manipulators. In this method to solve the path-planning task, the Transition-based Rapidly-exploring Random Tree (T-RRT) algorithm was applied, that requires a cost-function over the search space. The goal of this cost-function is to keep distance between the actual position and configurations that cause collisions. The paper presents two possible cost function generation methods. The first one is based on multidimensional Gauss functions and the second solution uses fuzzy function-approximation to determine the cost all over the configuration space. At last, the trajectory generator method will be introduced that calculates the desired shape of the function of joint variables and their derivatives. The algorithm is universal, the only restriction is that the manipulator and the environment are modelled by their bounding polyhedra. To demonstrate the presented approach, simulation were performed in MATLAB Simulink environment using the Mitsubishi RV-2F-Q robotic manipulator.

1 Introduction

Nowadays, automation is playing an increasingly important role in our daily lives, therefore is a growing need for the autonomous motion-planning of robotic manipulators.

Two types of path-planning methods can be defined, the offline and the online methods. In case of offline path-planning method, the calculation time does not mean a bottleneck, so more time can be spent on improving the found path. For example, in [1] an offline path-planning method is introduced, that can find an optimal collision-free path, where the optimality is based on the energy usage of the manipulator. These kind of methods can ensure the collision-freeness only with static obstacles or obstacles with known motions.

Online path-planning methods are used to avoid collisions with dynamic obstacles having unpredictable movements. These algorithms can be used in human-robot collaboration systems for example [2]. The environment of the robot can be observed with a depth camera. Besides, the image processing can be calculated with GPU, so the real-time calculation can be ensured [3]. In this paper, an offline path planning method will be described.

In our case, the minimization of cost-function is needed during the motion-planning algorithm, while avoiding configurations leading to collisions. There are numerous solutions in robotic applications to this problem.

Many of these are founded on the classical grid-based methods like A* or D* which can be used to find an optimal path limited by the resolution of the grid [4]. The main disadvantage of these algorithms is that the computation time grows exponentially with the number of dimensions. This effect is called as the "Curse of dimensionality".

Another type of path-planning methods are the sampling-based algorithms like the Rapidly-exploring Random Tree (RRT). These algorithms can be used effectively with higher number of dimensions as well, but the optimality of the solution is not guaranteed [5].

These problems can be solved by the Transition-based RRT (T-RRT) algorithm, that incorporates the advantages of the grid-based and the sampling-based methods [6]. Thus, it can be used to find a path that minimizes a cost-function defined in the high dimensional search space.

Heuristics can be used to determine the meaning of cost. For example, in case of a human-collaboration, to ensure the safe operation and the security, the cost has to be high near by the person [7].

Another method could be that, when the dynamics of the system is used to minimize the time or the performed work during the motion.

In this paper, the goal is to maximize the distance between the path of the manipulator and the configurations that cause collisions with the help of defined cost function. In this case the result of the offline path-planning algorithm can be used by a reactive mo-

*Corresponding Author: Dániel Szabó, szabo.daniel0604@gmail.com

tion planning method, that will have more space to avoid dynamic obstacles.

The paper is organized as follows. Section 2 introduces the base of the kinematics of robotic manipulators. In Section 3, the T-RRT algorithm is presented, that was used for path-planning in this work. Then the collision-detection method is described, based on investigation of the feasibility of inequalities of bounding polyhedra, that are describing the manipulator and the obstacles (Section 4). Then, Section 5 describes two cost function approximation algorithms with Gauss-functions and with fuzzy function-approximation. The latter is more detailed. Subsection 5.5 describes the method that is used in this work for cost function evaluation. The trajectory generation method is explained in Section 6. The simulation results are presented in Section 7. Finally, Section 8 contains conclusions and possible developments in the future.

2 Robotic manipulators

Robotic manipulators are usually modeled as open kinematic chains. A manipulator contains links, which are connected by joints. Two basic types of joints exist: Revolute joints are used to apply relative rotation, and prismatic joints allow linear motion between two adjacent links. The relative movement between adjacent links is given by the joint variables q_i. In case of revolute joints, q_i gives the angular displacement, while for prismatic joints, q_i equals the amount of linear motion. The actual configuration q of the robot is defined by the vector containing actual q_i values of all joints.

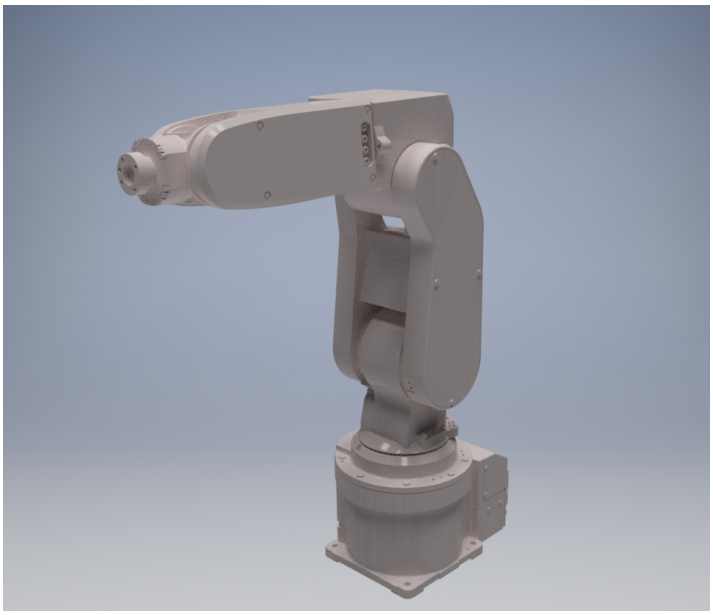


Figure 1: Mitsubishi RV-2F-Q manipulator (used in the simulations)

An end effector is usually attached to the last segment of the manipulator. To determine the position and the orientation of the end effector in the workspace of the robot, the equations of forward kinematics can be used. A homogeneous transformation matrix belongs to each robotic joints. The position and orientation of the coordinate frame attached to the end effector (relative to the base frame) can be determined by multiplying these matrices.

The main parameters of a robotic manipulator can be defined by the Denavit-Hartenberg parameters [8]. Homogeneous transformation matrices can be calculated from these parameters using actual joint variables.

The task of the path-planning is to find a path, which moves the robot to a desired goal configuration. (This goal configuration is typically given by the desired orientation and position of the end-effector at the end of the motion.) The path-planning problem can be defined in the Euclidean space, but usually, the planning is performed in the configuration space $\mathcal{C}$ of robot. The configuration space is spanned by the admissible configuration vectors (q), i.e. joint variables (q_i) are used as coordinates. For example the configuration space of the Mitsubishi RV-2F-Q manipulator with six degrees of freedom is a six-dimensional space. This robot has six revolute joints. Its 3D model is depicted in Fig. 1. The simulations presented in Section 7 were performed using this manipulator.

3 The Path-planning Algorithm

The traditional RRT algorithm is a stochastic sampling path-planning algorithm. In this method a search tree is grown from the initial q_{init} point to the q_{goal} goal configuration. In a given iteration, a randomly selected configuration is inserted to the tree [9].

In these days, the planning algorithms like the RRT method are very popular, due to their efficiency in exploration of the search space. There exist several improvements of the traditional method, for example the RRT* described in [10], which checks the available corrections in every iteration cycle, so it is able to find an optimal path to the goal point. The method presented in [11] called RRT$^+$ can be used in very high dimensional spaces, such as the configuration space of hyper-redundant manipulators. In this work, the T-RRT algorithm is used, that is able to find a suboptimal solution defined by a cost function in high dimensional configuration spaces as well. This solution is called as high quality path in [6]. In our case the goal of cost function is to minimize the risk of collision.

Algorithm 1: Rapidly-exploring Random Tree (RRT)

Input: the configuration space $\mathcal{C}$
the root q_{init} and the goal q_{goal}
Output: the tree $\mathcal{T}$

1: $\mathcal{T} \leftarrow$ InitTree(q_{init})
2: **while not** StopCondition($\mathcal{T}, q_{goal}$) **do**
3: $q_{rand} \leftarrow$ SampleConf($\mathcal{C}$)
4: $q_{near} \leftarrow$ NearestNeighbor($q_{rand}, \mathcal{T}$)
5: $q_{new} \leftarrow$ Extend($\mathcal{T}, q_{rand}, q_{near}$)
6: **if** NoCollision(q_{new}) **then**
7: AddNewNode($\mathcal{T}, q_{new}$)
8: AddNewEdge($\mathcal{T}, q_{near}, q_{new}$)
9: **end if**
10: **end while**
11: **return** $\mathcal{T}$

In Algorithm 1 the traditional RRT algorithm is described. The inputs of the algorithm is the configuration space, where the path-planning method is evaluated, the initial and the goal configuration of the path-planning. The output of the method is the search tree. At first, the search tree is initialized with the initial configuration as its root node.

After that a uniformly distributed random configuration q_{rand} is generated in function SampleConf.

This configuration is used by the NearestNeighbor function to determine the nearest configuration (called q_{near}) to it in the search tree. In this method, distance from the edges is not examined, due to the higher computational demand, it is determined only from the previously added nodes.

A new point q_{new} is created in Extend, which is in the same direction as q_{rand} to q_{near}, but the distance between q_{new} and q_{near} is maximized with the parameter δ. In this method, the collision-freeness of the created q_{new} configuration is not examined.

The function NoCollision returns true, if the q_{new} configuration does not lead to a collision. In this case, a new node and edge will be added to the tree.

The algorithm is succeeded, if the distance between the last inserted point and the goal configuration is less than a given parameter, or failed, if a maximum number of iterations has been reached. In the function StopCondition are these conditions inspected.

Algorithm 2 describes the T-RRT method. As it can be seen, the method is based on the traditional RRT algorithm described in Algorithm 1. The main differences are the TransitionTest and MinExpandControl functions (see Subsections 3.1-3.2), which are the conditions to extend the search tree with a new node and edge.

The cost of the found path can be evaluated by several methods, e.g. maximal cost, average cost or total cost along the path. However it will lead to a path, that avoids the areas of search space with higher costs, if the performed work is minimized along the path, as it is introduced in [6]. In this case, the positive variations of the cost function will be minimized.

Algorithm 2: Transition-based RRT

Input: the configuration space $\mathcal{C}$
the cost function $c : \mathcal{C} \mapsto \mathbb{R}_+^*$
the root q_{init} and the goal q_{goal}
Output: the tree $\mathcal{T}$

$\mathcal{T} \leftarrow$ InitTree(q_{init})
while not StopCondition($\mathcal{T}, q_{goal}$) **do**
  $q_{rand} \leftarrow$ SampleConf($\mathcal{C}$)
  $q_{near} \leftarrow$ NearestNeighbor($q_{rand}, \mathcal{T}$)
  $q_{new} \leftarrow$ Extend($\mathcal{T}, q_{rand}, q_{near}$)
  if $q_{new} \neq NULL$ **then**
    $d_{near-new} \leftarrow$ Distance(q_{near}, q_{new})
    if TransitionTest($c(q_{near}), c(q_{new}), d_{near-new}$)
    and MinExpandControl($\mathcal{T}, q_{near}, q_{rand}$)
    and NoCollision(q_{new}) **then**
      AddNewNode($\mathcal{T}, q_{new}$)
      AddNewEdge($\mathcal{T}, q_{near}, q_{new}$)
    end if
  end if
end while
return $\mathcal{T}$

3.1 Transition-test

Algorithm 3 describes the TransitionTest function. In the first step, the number, how many times the transition-test failed previously, is queried by the GetCurrentNFail function.

Thereafter if the cost in q_{new} point is higher than a c_{max} parameter, then the test will not be accepted, so the search tree will not be extended. In case of negative variation of cost function the transition test will return true.

After that in case of positive slope of the cost, the transition test will be successful in accordance with a probability based on the Metropolis criterion [12], defined as

$$p_{ij} = \exp(-\frac{\Delta c_{ij}}{K \cdot T}) \tag{1}$$

where $\Delta c_{ij} = \frac{c_j - c_i}{d_{ij}}$ is the variation of the cost, K is a normalization constant, it can be taken as the average cost of the query configurations.

The difficulty of success for a given transition is defined by the temperature parameter denoted by T. The goal of this parameter is to minimize the amount of positive slopes in the cost function along the path, but it allows to explore the whole reachable configuration space by its adaptive tuning. In case of higher temperature the acceptance of transition test is more probable for a transition with higher positive slope. To accept only very low positive variations at the start, T is initialized with a very low value. This functions similar to the simulated annealing algorithm.

The function Rand(0,1) gives a random number between 0 and 1 based on uniform distribution.

The $\alpha > 1$ variable is used to change the temperature adaptively. If the transition was successful then the temperature is decreased. And if the test failed $nFail_{max}$ times then the value of T is increased by multiplying with α. Thus if the $nFail_{max}$ parameter is high then the path will be more likely to go into the valleys, however finding a path over a region with high cost will be more difficult.

Algorithm 3: TransitionTest(c_i, c_j, d_{ij})

$nFail$ = GetCurrentNFail();
if $c_j > c_{max}$ **then**
  return False
end if
if $c_j < c_i$ **then**
  return True
end if
$p = \exp(\frac{-(c_j - c_i)/d_{ij}}{K \cdot T})$
if Rand(0, 1) $< p$ **then**
  $T = T/\alpha$
  $nFail = 0$
  return True
else
  if $nFail > nFail_{max}$ **then**
    $T = T \cdot \alpha$
    $nFail = 0$
  else
    $nFail = nFail + 1$
  end if
  return False
end if

3.2 Minimal Expansion Control

The transition-test can slow the exploration rate of high cost hills and the new configurations only refine the already known regions.

By specifying a minimum level of exploration, this can be prevented. Accordingly, a new node close to the closest search tree configuration will only be accepted if the refining node ratio is smaller than the ρ parameter. The minimal exploration control method will accept the exploring nodes automatically.

This function is introduced in Algorithm 4. The function UpdateNbNodeTree increases the number of nodes, while the UpdateNbRefineNodeTree function updates the number of refining nodes in the search tree.

The Euclidean distance between the q_{rand} (not the extended q_{new}) and the q_{near} configurations is used to determine the type of a given point. A point is an exploring node if the distance is high, otherwise it will be a refining node.

The path-planning will discover the entire configuration space very effectively with this complement, while trying to stay in the valleys of the cost function owing to the transition-test.

Algorithm 4: MinExpandControl($\mathcal{T}, q_{near}, q_{rand}$)

```
if Distance(q_near, q_rand) > δ then
    UpdateNbNodeTree(T)
    return True
else
    if NbRefineNodeTree(T+1) / NbNodeTree(T+1) > ρ then
        return False
    else
        UpdateNbRefineNodeTree(T)
        UpdateNbNodeTree(T)
        return True
    end if
end if
```

4 Collision detection

The effectiveness of a path planning method depends highly on the applied collision detection function, since it is called often (see Step 6 of Algorithm 1 or Step 8 of Algorithm 2) and it has to ensure that only collision-free configurations are used for planning. To calculate in real time, which robot configurations are in collisions, is a difficult problem. The task is not only to determine if there is any collision between the robot and the surrounding objects, but also the self-collision configurations of the robot has to be avoided.

Moreover, the goal could be not only to plan a collision-free path, but also to get a time-optimal motion. A solution for such a problem is presented in [1], but the computational demand of that algorithm is high, due to the solution of the optimization problem. However, the collision detection algorithm used by that method can be applied by other applications as well. The algorithm presented in this paper is also based on this collision detection.

4.1 Collision detection with polyhedrons

The basic idea of this collision detection method is that the geometry of the robot and the obstacles is approximated such that they are represented by unions of convex polyhedrons. Fig. 2 and Fig. 3 illustrate how the bounding polyhedrons can be determined around the robot and the obstacles. For the sake of simplicity, it is supposed that the workspace of the robot contains only one obstacle. The presented collision detection method can be easily extended if more obstacles are present.

Let P denote the union of polyhedrons that represent the robot:

$$P = \bigcup_{i=1}^{n_P} P^{(i)} = \bigcup_{i=1}^{n_P} \{y \in \mathbb{R}^3 | A^{(i)} y \leq b^{(i)}\} \tag{2}$$

where n_P shows the number of polyhedrons in P. A polyhedron is bounded by faces. A face $P^{(i)}$ is defined by $A^{(i)}$ and $b^{(i)}$ parameters and p_i is the number of faces. Then $A^{(i)} \in \mathbb{R}^{p_i \times 3}$ and $b^{(i)} \in \mathbb{R}^{p_i}$ for $i = 1, \ldots, n_p$.

Similarly, an obstacle Q can also be described by the union of polyhedrons:

$$Q = \bigcup_{j=1}^{n_Q} Q^{(j)} = \bigcup_{j=1}^{n_Q} \{y \in \mathbb{R}^3 | C^{(j)} y \leq d^{(j)}\} \tag{3}$$

where n_Q denotes the number of polyhedrons in Q. q_j shows the number of faces in $Q^{(j)}$, hence $C^{(j)} \in \mathbb{R}^{q_j \times 3}$ and $d^{(j)} \in \mathbb{R}^{q_j}$ for $j = 1, \ldots, n_Q$.

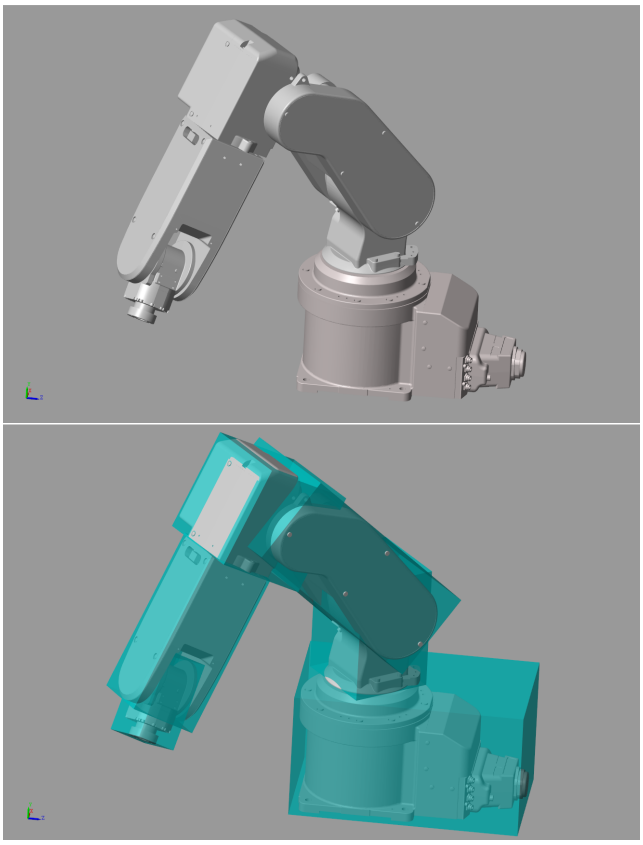


Figure 2: Bounding boxes of each segments were defined as the bounding polyhedrons of the manipulator

To avoid self-collisions, pairs of polyhedrons $P^{(i)}$ have to be checked, whether they collide with each-other. There exist such

pairs of $P^{(i)}$ where the collision detection is not necessary, since they cannot collide physically (e.g. adjacent segments). Therefore, the set of pairs $(k, l), k \neq l$ has to be defined, such that k and l denote two polyhedrons $(P^{(k)}, P^{(l)})$ of the robot which could collide. This set is denoted by I. If $(k, l) \in I$, the collision check has to be performed for the corresponding two polyhedrons.

Suppose, that P is already calculated for a given configuration and Q is determined as well, then the configuration is collision-free if

$$P^{(i)} \cap Q^{(j)} = \emptyset \quad \wedge \quad P^{(k)} \cap P^{(l)} = \emptyset \tag{4}$$

$\forall i = 1, \ldots, n_P$ and $\forall j = 1, \ldots, n_Q$ and $\forall (k, l) \in I$.

One polyhedron of the robot $P^{(i)}$ and one of the obstacle $Q^{(j)}$ do not collide with each-other if the union of their system of inequalities has no solution. Formally, there exists no such $y^{(i,j)} \in \mathbb{R}^3$ point where

$$\begin{pmatrix} A^{(i)} \\ C^{(j)} \end{pmatrix} y^{(i,j)} \leq \begin{pmatrix} b^{(i)} \\ d^{(j)} \end{pmatrix} \tag{5}$$

Similar inequalities can be used to check self-collision.

To determine, whether an inequality similar to (5) has no solution, Farkas's lemma can be applied. The lemma says that there exists no solution for the linear system if and only if a vector $w^{(i,j)} \in \mathbb{R}^{(p_i+q_j)}$ can be found such that

$$w^{(i,j)} \geq 0 \quad \text{and} \quad \begin{pmatrix} A^{(i)} \\ C^{(j)} \end{pmatrix}^{\top} w^{(i,j)} = 0 \quad \text{and} \tag{6}$$

$$\begin{pmatrix} b^{(i)} \\ d^{(j)} \end{pmatrix}^{\top} w^{(i,j)} < 0$$

Consequently, if given is a configuration q and a $w^{(i,j)}$ solution can be found for (6), then the q configuration can be marked as collision-free.

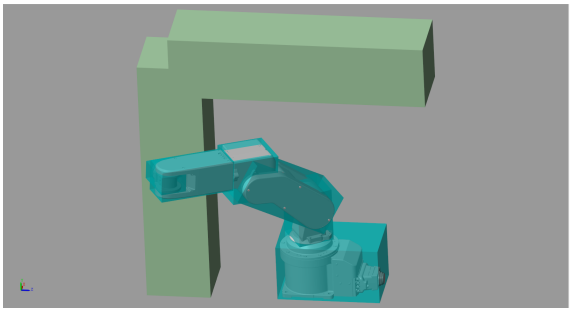


Figure 3: Bounding polyhedrons around the robot and obstacles (as used in the simulation)

4.2 Construction of polyhedrons

The face parameters $(A^{(i)}, b^{(i)})$ for some robot-polyhedron $P^{(i)}$ depend on the actual robot configuration q. This subsection describes how to calculate them.

First, $A_r^{(i)}$ and $b_r^{(i)}$ have to be determined for each segment of the robot in a reference coordinate system that is attached to the same segment. For example, the CAD model of the manipulator (e.g. shown in Fig. 1) can be used to measure the corresponding parameters in this frame. (For obstacles, the same process has to be performed.)

Then, a transformation is required to get these parameters in the base frame. Denote the resultant transformation matrix by $T^{(i)}(q)$. $T^{(i)}(q)$ can be used for the transformation from the base coordinate system to the reference frame of the i_{th} segment, if the actual configuration of the robot is q.

The parameters in the base frame can be determined as follows

$$\begin{bmatrix} A^{(i)} & b^{(i)} \end{bmatrix} = \begin{bmatrix} A_r^{(i)} & b_r^{(i)} \end{bmatrix} \begin{bmatrix} T^{(i)}(q) \end{bmatrix}^{-1} \tag{7}$$

During the collision-check algorithm, these parameters have to be used.

5 Cost-function approximation

The goal of cost-function is to keep distance from configurations leading to a collision. The approximation of this function is necessary, to decrease the computational demand of the algorithm. This cost-function is used by the T-RRT path-planning method, introduced in Section 3.

5.1 Cost-funcion approximation with Gauss-functions

At first, the approximation of cost-function was evaluated by taking enough sample of the configuration space and Gauss-functions were summed in every sampling point. The centers of Gauss-functions are the configurations causing a collision. There are several disadvantages of this method.

As it can be seen in the Fig. 4 the result is very hilly and a maximal value of the cost function can not be predefined, so the c_{max} parameter of the TransitionTest function in Algorithm 3 is difficult to determine.

In addition, the cost function has to be stored in every sampling point, moreover the interpolation is slow at any other point.

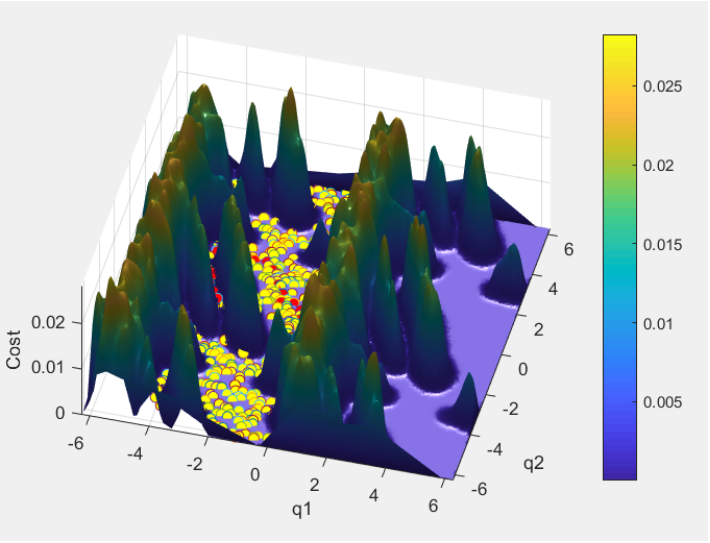


Figure 4: Cost-function approximation with Gauss-functions with the search tree found by the T-RRT algorithm

5.2 The base of fuzzy approximation

Another solution could be the fuzzy approximator systems.

Fuzzy systems have several uses in many areas such as decision making [13], system identification [14], control theory [15] etc. In addition, the approximation of nonlinear functions is possible as well, if enough teaching points are available. Teaching points are defined as $y_j = f(x_j)$, where f is the function that needs to be approximated, x_j is the position of the given teaching point and y_j is the value of f in x_j.

There are several advantages of fuzzy estimator systems. Depending on which type of inference method is used, their rules can be tuned adaptively and the approximation time is one of the benefits of these systems as well.

Singleton fuzzyfication, center average defuzzyfication, product inference and zero-order Sugeno model fuzzy systems with Gaussian membership functions are therefore the most widely used approaches [16].

Thus the fuzzy estimation of a nonlinear function is defined as

$$\hat{f}(x) = \frac{\sum_{l=0}^{M} \bar{y}^l \prod_{i=1}^{n} \exp(-(\frac{x_i - \bar{x}_i}{\sigma_i^l})^2)}{\sum_{l=0}^{M} \prod_{i=1}^{n} \exp(-(\frac{x_i - \bar{x}_i}{\sigma_i^l})^2)} \tag{8}$$

where n is the number of dimensions in the configuration space, $\bar{y}^l$ is the output in the l_{th} rule of the fuzzy inference system, $\bar{x}_i$ is the location of the maximum value of the Gaussian membership function and with σ_i^l the width of this function can be modified.

The adaptive tuning of the fuzzy rules is feasible by modifying the $\bar{y}^l$, $\bar{x}_i$ and σ_i^l values.

According to the Universal approximation theorem described in [16], for any given real continuous function f on the compact set $U \in \mathbb{R}^n$ and arbitrary $\epsilon > 0$ there exists $\hat{f}$ in the form given in (8) such that

$$\sup_{x \in U} |f(x) - \hat{f}(x)| < \epsilon \tag{9}$$

Henceforth, two adaptive tuning method for fuzzy estimator systems will be introduced. In which bounded domain intervals of input vector x and output y are assumed.

5.3 Generating fuzzy rules by learning from examples

Details of the algorithm can be found in [17].

The input and output domain is divided into intervals, in each Gaussian membership functions are defined.

The possible fuzzy relations are determined exactly by the number of input membership functions. For example, in case of input variables x_1 and x_2, if the number of membership functions are N_1 and N_2 then the number of possible rules is $N_1 \cdot N_2$. For each teaching point the relation and output membership function has to be selected that mostly fits the given teaching point.

In k_{th} teaching step a generic form of fuzzy rules can be defined:

$$R^k : \text{IF } x_1 \text{ IS } F_1^k \text{ AND } x_2 \text{ IS } F_2^k \text{ AND } \ldots \text{ AND } x_n \text{ IS } F_n^k$$

$$\text{THEN } y \text{ IS } G^k$$

where F_i^k is one of the input variable membership functions and G^k is one of the output variable membership functions.

In the k_{th} step a given rule will be chosen, if it maximizes the following expression:

$$D(R^k) = \prod_{i=1}^{n} (\mu_{F_i^k}(x_i^k)) \cdot \mu_{G^k}(y^k) \tag{10}$$

where μ is the firing strength of the given membership function for the actual value of the teaching point.

There exist more possible solutions to that problem, if a new relation has to be added with the same condition part, but a different inference part as an already added relation. In this case, the relation with higher value can be selected by evaluating (10), or a weighted value of both inference parts can be specified.

The main disadvantage of this technique is that (10) has to be evaluated $N = N_1 \cdot N_2 \cdot \ldots \cdot N_n$ times for every teaching points (where n is the number of dimensions of the space) to find the maximum value. As a consequence, with the number of dimensions the computational time grows exponentially.

5.4 Nearest neighborhood clustering

In this algorithm, teaching points are clustered depending on their position in the configuration space. A given cluster can be described by three parameter, the center of it x_0^l, the number of teaching points in it B^l and the sum of values of these points A^l [18].

A new point x_j ($y_j = f(x_j)$) is inserted to the cluster l, if it is closer to x_0^l than a given r radius. After that the other cluster parameters A^l and B^l will be updated. If the distance of the teaching point x_j is too high from all of the clusters then a new cluster will be created as $\{x_0^l = x_j; A^l = y_j; B^l = 1\}$. M represents the number of clusters that have been defined during the teaching process.

Therefore, the estimation of the nonlinear function can be defined as

$$\hat{f}(x) = \frac{\sum_{l=0}^{M} A^l \exp(-(\frac{x - \bar{x}_0^l}{\sigma^l})^2)}{\sum_{l=0}^{M} B^l \exp(-(\frac{x - \bar{x}_0^l}{\sigma^l})^2)} \tag{11}$$

5.5 Cost-function evaluation

Eventually, the cost-function for the T-RRT method was approximated with the Nearest neighborhood clustering algorithm, because it has a higher computational efficiency.

To determine the position of the teaching points, random sampling of the configuration space is needed, owing to the high number of dimensions.

The collision detection method described in Section 4 is evaluated in all of teaching points. If the x_j configuration causes collision, then one will be assigned as the value of the cost-function y_j in x_j, otherwise zero.

After that the rules of the fuzzy approximator are tuned with the Nearest neighborhood clustering algorithm. This method can be used for arbitrary robotic manipulators. An example for the cost function and path can be seen in Fig. 5 in case of 2-DoF manipulator (since it is not possible to depict the cost function in higher dimensions).

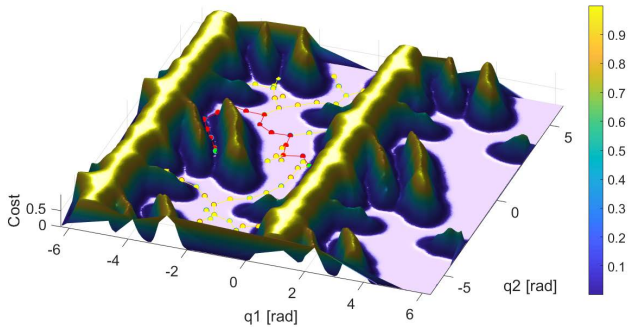


Figure 5: T-RRT method is able to find high quality path based on a cost-function. The cost-function is evaluated for a two degrees of freedom manipulator with the Nearest Neighborhood Clustering algorithm.

Table 1 shows the comparison of approximation with Gauss-function and the Nearest neighborhood clustering algorithm. As can be seen, both the training time and the evaluation time for one configuration in the case of Gauss-function approximation grow very fast by increasing the number of teaching points. In the case of Nearest neighborhood clustering, these values increased to a lesser extent. By using clusters, the evaluation time of the fuzzy function-approximation method did not increase significantly, which is ideal for applying the method in a path-planning algorithm.

Table 1: Comparison of fuzzy function-approximation and approximation with Gauss-functions

Num. of teaching points	Nearest Neighborhood clustering			Gauss	
	Training time	Number of clusters	Eval. time	Training time	Eval. time
1000	0.0481s	714	22μs	0.0402s	7.8μs
5000	0.207s	1649	45μs	0.3027s	28.8μs
10000	0.4554s	1974	52μs	1.4151s	58.2μs
50000	3.0962s	2431	65μs	22.88s	137.9μs
100000	7.045s	2521	76μs	155.26s	1632μs

6 Trajectory generation

The already presented methods can be used to find points in the configuration space, that by moving the robot between these points, the possible collisions can be avoided.

6.1 Trajectory generation for scalar values

In the following method y denotes one coordinate of the q configuration, in other words, one joint variable of the robot.

Let $\{y_k\}$ be the set of points founded by the path-planning algorithm and $\{t_k\}$ denotes the desired absolute time of reaching these points [19].

$$\begin{aligned} \{y_k\} &= \{\dots, A, B, C, D, \dots\}, \\ \{t_k\} &= \{\dots, t_A, t_B, t_C, t_D, \dots\} \end{aligned} \tag{12}$$

Due to the limits of torques, boundaries have to be defined for the acceleration of joints: $|\ddot{y}|_{max}$.

At first, $y(t)$ trajectory can be defined as a series of linear functions between adjacent points. In this case, the first derivative $\dot{y}(t)$ is constant between two points and it steps to another constant value in a given point. The acceleration $\ddot{y}(t)$ is zero in every point except the points where the slope of $y(t)$ trajectory is changing. In these points, the second derivative contains Dirac-delta functions, which leads to an unbounded control value, so this solution is not applicable. To overcome this problem, the acceleration of joints have to be changed continuously.

To reach the next point two kinds of phases can be defined: traveling phase and acceleration phase. In the former, we use a constant velocity and zero acceleration, while in the latter a continuous acceleration function has to be applied, to reach the desired constant velocity of the next traveling phase. The quadratic function can be chosen as the form of the function in the acceleration phase (Fig. 6).

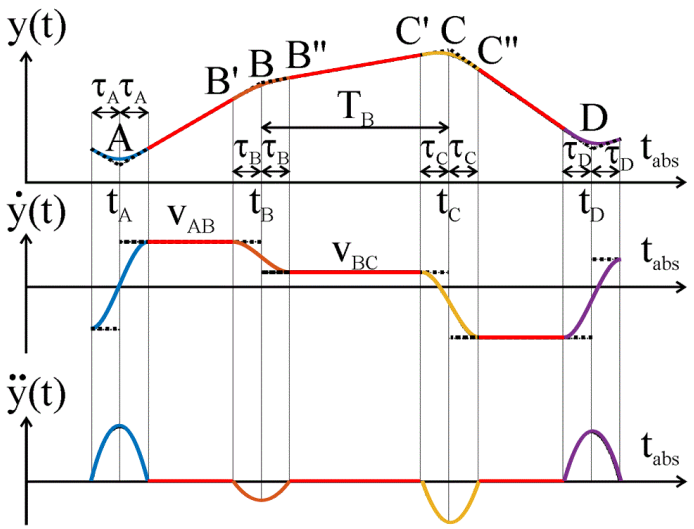


Figure 6: To ensure the continuity of the acceleration function, quadratic function can be applied in the acceleration phase. The form of the trajectories can be seen in the figure.

Let B and C be two adjacent points, while B' and C' are the beginning, B'' and C'' are the end of their acceleration phase. The relative time t can be defined as

$$t = t_{abs} - t_B \in [-\tau_B, T_B], \tag{13}$$

where t_B is the absolute time of reaching the point B, τ_B is the half of duration of acceleration phase in B and T_B denotes the time-interval of reaching point C from point B. This can be seen in Fig 7. The time-intervals between configurations can be determined by the boundaries of joint-velocities, for example $T_B = \frac{C-B}{|\dot{y}|_{max}}$.

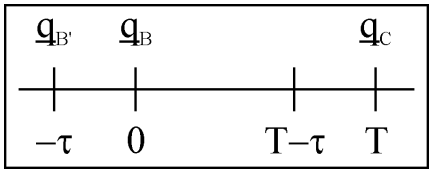


Figure 7: Relative time can be defined, where the zero value is in the q_B configuration. The zero point will be shifted with T, when the $T - \tau$ relative time is reached.

By integrating the quadratic acceleration during the acceleration phase, the velocity and position can be determined:

$$\begin{aligned} B'B'': \quad \ddot{y}(t) &= a_0 t^2 + a_1 t + a_2 \\ \dot{y}(t) &= a_0 \frac{t^3}{3} + a_1 \frac{t^2}{2} + a_2 t + a_3 \\ y(t) &= a_0 \frac{t^4}{12} + a_1 \frac{t^3}{6} + a_2 \frac{t^2}{2} + a_3 t + a_4 \end{aligned} \tag{14}$$

To determine the $a_0 \dots a_4$ parameters of the trajectory of position, five independent conditions have to be defined:

$$\begin{aligned} \ddot{y}(-\tau_B) &= 0 \\ \ddot{y}(\tau_B) &= 0 \\ \dot{y}(-\tau_B) &= v_{B'B} = \frac{B - B'}{\tau_B} \\ \dot{y}(\tau_B) &= v_{BC} = \frac{C - B}{T_B} \\ y(\tau_B) &= B + v_{BC} \cdot \tau_B \end{aligned} \tag{15}$$

The following linear equation system can be defined:

$$\begin{bmatrix} \ddot{y}(-\tau_B) \\ \ddot{y}(\tau_B) \\ \dot{y}(-\tau_B) \\ \dot{y}(\tau_B) \\ y(\tau_B) \end{bmatrix} = \begin{bmatrix} \tau_B^2 & -\tau_B & 1 & 0 & 0 \\ \tau_B^2 & \tau_B & 1 & 0 & 0 \\ -\frac{\tau_B^3}{3} & \frac{\tau_B^2}{2} & -\tau_B & 1 & 0 \\ \frac{\tau_B^3}{3} & \frac{\tau_B^2}{2} & \tau_B & 1 & 0 \\ \frac{\tau_B^4}{12} & \frac{\tau_B^3}{6} & \frac{\tau_B^2}{2} & \tau_B & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \tag{16}$$

$$\mathbf{y} = \mathbf{C} \cdot \mathbf{a}$$

The solution of this linear equation system can be calculated as follows:

$$\mathbf{a} = \mathbf{C}^{-1} \cdot \mathbf{y} \tag{17}$$

In the traveling phase ($t \in [\tau_B, T_B - \tau_C)$), the form of the path of the scalar value will be a linear function:

$$y(t) = B + v_{BC} \cdot t \tag{18}$$

6.2 Trajectory generation for joint variables

The presented method for scalar values can be used for each joint variable separately, by taking into account the difference between the boundaries of velocities ($|\dot{q}_i|_{max}$) and accelerations ($|\ddot{q}_i|_{max}$) of some joints.

In addition, the minimal time needed to reach q_{Ci} joint value from q_{Bi} can be different as well, so just that joint have to be moved in maximal speed, that needs the most time to reach the next configuration.

The same τ parameter can be used for each joint variable and it is calculated as follows [19]:

$$\tau = \max_i \frac{3}{2} \frac{|\dot{q}_i|_{max}}{|\ddot{q}_i|_{max}} \tag{19}$$

The trajectory generation method can be seen in the Algorithm 5. The input of the method denoted by Q is the series of configurations contained by the found path.

Algorithm 5: TrajectoryGeneration(Q, τ)

```
q_C = GetFirstConfiguration(Q)
T = τ
while Size(Q) > 0 do
    q_B' = q(T − τ)
    q_B = q_C
    q_C = GetFirstConfiguration(Q)
    T_i = |q_Ci − q_Bi| / |q̇_i|_max
    T = max_i{max{T_i}, 2τ}
    v_AB = (q_B − q_B') / τ
    v_BC = (q_C − q_B) / T
    a = C^{-1} · y
    t = −τ
    while t < T − τ do
        q(t) = CalculateTrajectory(a, t)
        t = t + Δ
    end while
end while
return q(t)
```

The GetFirstConfiguration function gives the first element of the configuration queue and this element will be removed from the queue as well.

At first the variable T is initialized with τ calculated by (19), so in the first iteration the $q_{B'}$ variable will be the root configuration.

The equations in (14) are evaluated in the CalculateTrajectory function for each joint variable, consequently the dimension of the q variable is equal to the dimension of the configuration space.

Trajectories calculated with this method can be seen in Fig. 9.

7 Simulation results

The comparison of several path-planning algorithms can be seen in Table 2 in case of two degrees of freedom robotic manipulator. The result of planning with the T-RRT method can be seen in Fig. 8. The RRT* algorithm is a variant of traditional RRT method, that refines the search tree in every iteration to find the actual shortest path to the goal configuration [10]. In case of multi-directional RRT method an other search tree is started from the goal configuration [5]. The calculation time can be reduced by this method.

Table 2: Comparison of path-planning algorithms

Algorithm	Calculation time [s]	Length	Total cost	Performed work
RRT	0.039	10.54	0.415	0.896
Multi-directional RRT	0.019	10.54	0.396	0.822
Multi-directional RRT*	3.34	6.92	0.171	0.449
T-RRT$_g$	0.55	10.21	0.028	0.121
T-RRT$_t$	13.466	10.91	0.01	0.062

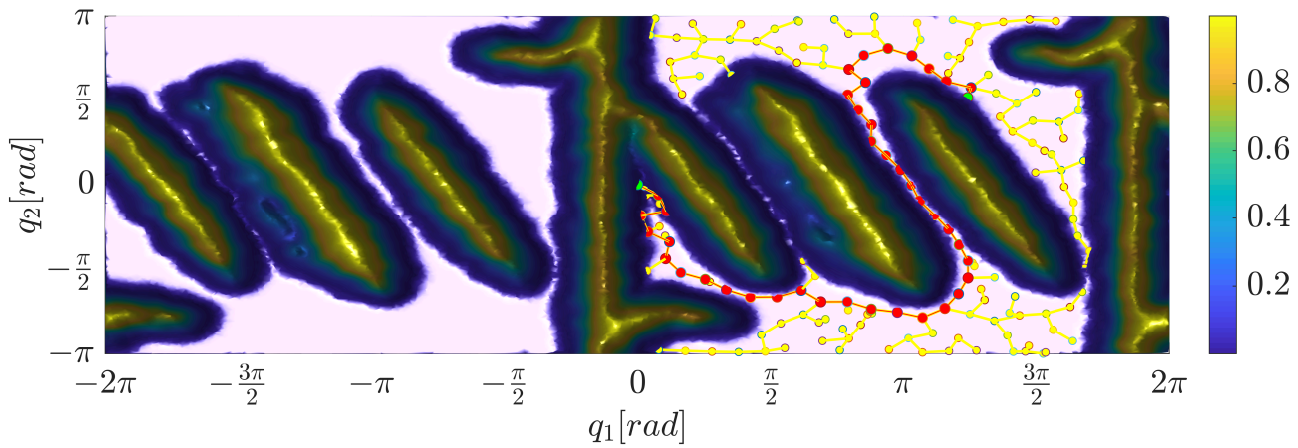


Figure 8: The result of path-planning for two degrees of freedom manipulator with the tempered version of T-RRT

Two parameter settings are examined for the T-RRT method, the first one (T-RRT$_g$) is a greedy version with $nFail_{max} = 10$, in the other case (T-RRT$_t$) $nFail_{max} = 100$, which will lead to a tempered version of T-RRT. The computation time in the first case is lower, but the found path has a lower quality as well. The other parameters were: $T_{init} = 10^{-5}$, $\alpha = 1.5$, $\rho = 0.2$, $c_{max} = 0.4$, $\delta = 0.3$. The parameters were chosen empirically.

As it can be seen, the path of the T-RRT algorithm is not equal with the shortest collision-free path.

To demonstrate the presented method, simulations were performed using the Mitsubishi RV-2F-Q manipulator as well. The task of the robot was to move to the other side of an obstacle, which is modeled by two polyhedrons. For implementation and simulation, MATLAB Simulink and Simscape Multibody toolbox were used. The running times were measured on a computer with Intel Core i7-7700HQ Processor.

For the cost function evaluation, the nearest neighborhood clustering method was applied. The parameters were empirically selected as follows: number of teaching points is $n = 20000$, radius of the cluster is $r = 1$ *rad* and deviation of the Guassian membership functions is $\sigma = 0.5$. To determine the approximated cost function, $t \approx 120s$ computational time was required, what is appropriate, since these calculations have to be performed only once for a given robot and workspace.

The parameters of the T-RRT method were selected as $T_{init} = 0.01$, $\alpha = 5$, $nFail_{Max} = 10$, $\rho = 0.1$, $c_{max} = 0.8$ and $\delta = 0.4$. To get a solution for a path-planning problem, the average calculation time was $t \approx 7.3s$.

One solution is presented in Fig. 9–Fig. 10. Fig. 9 depicts the trajectories of the joint variables. The whole movement can be seen in a video: `https://youtu.be/t6GO8LKG5js`. Some key configurations of the robot are also presented in Fig. 10. It can be seen, that the presented algorithm was able to plan such a motion, where colliding configurations can be avoided.

based on the T-RRT method, which is able to calculate the reference motion for the robot, such that collision avoidance is guaranteed and suboptimal path is selected according to some cost-function.

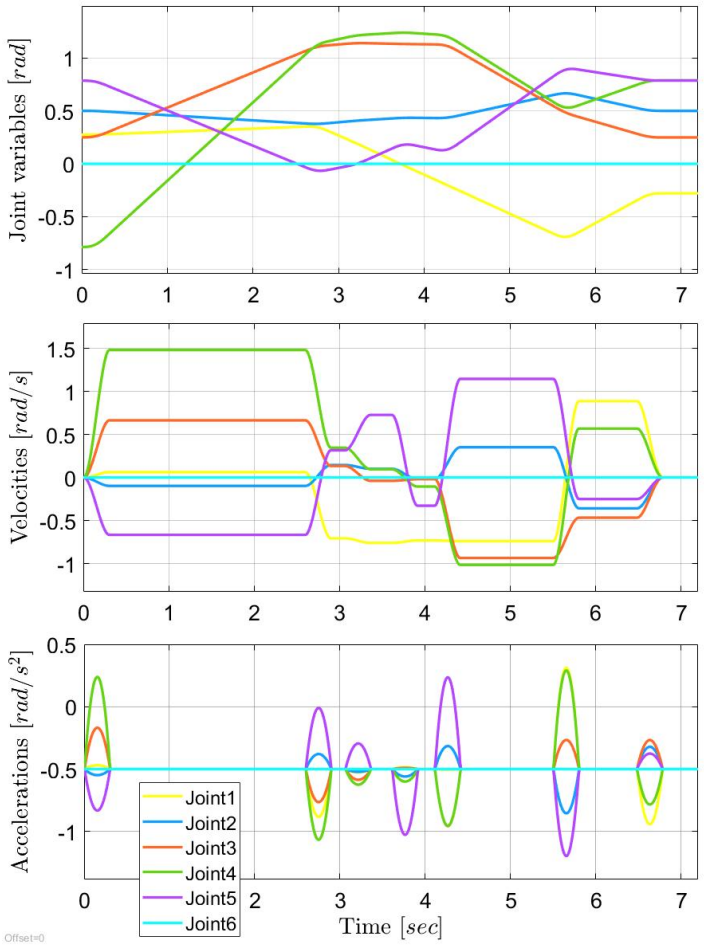


Figure 9: Trajectories of the joint variables

8 Conclusion

The method presented in this paper can be used to solve the offline path-planning problem for robotic manipulators. The algorithm is

In this work such a cost-function was selected, which can ensure, that the distance between the path of the robot and the colliding configurations is large. Instead of the time-consuming determination of

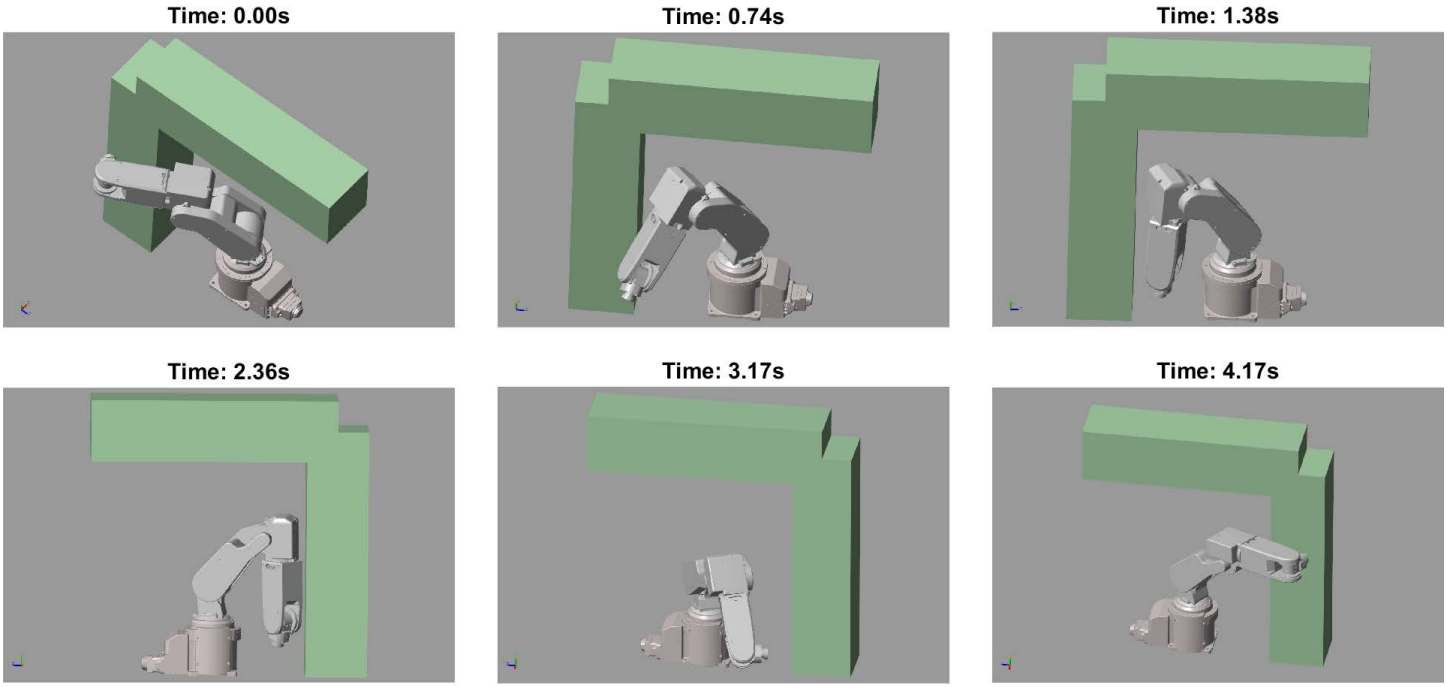


Figure 10: Robot configurations along the collision-free path

the exact cost values, a nonlinear-fuzzy approximation algorithm was suggested which uses the Nearest neighborhood clustering method.

For the collision checking of robot configurations, another approximation was applied. The segments of the robot and the obstacles can be modeled by sets of bounding polyhedrons. The collision of these polyhedrons was checked by the feasibility analysis of inequality systems.

Trajectory planning was also presented in the paper. To produce an input for a trajectory tracking and control method, time functions were determined for the joint variables and their derivatives using known bounds of the joint velocities and accelerations.

In the sequel, the goal is to develop an effective online motion-planning method, which can be used to plan a collision-free motion in such a workspace which contains obstacles with unpredictable movements as well. First, the reference motion can be determined by the presented T-RRT based method for the static environment, and additionally, a reactive planning method (similar to [20]) can be applied to ensure the avoidance of dynamic obstacles as well.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment The research reported in this paper was supported by the Higher Education Excellence Program in the frame of Artificial Intelligence research area of Budapest University of Technology and Economics (BME FIKP-MI/SC).

References

[1] Chantal Landry, Matthias Gerdts, René Henrion, and Dietmar Hömberg. Path-planning with collision avoidance in automotive industry. In *IFIP Conference on System Modeling and Optimization*, pages 102–111. Springer, 2011.

[2] Fabrizio Flacco, Torsten Kröger, Alessandro De Luca, and Oussama Khatib. A depth space approach to human-robot collision avoidance. In *2012 IEEE International Conference on Robotics and Automation*, pages 338–345. IEEE, 2012.

[3] Massimo Cefalo, Emanuele Magrini, and Giuseppe Oriolo. Parallel collision check for sensor based real-time motion planning. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1936–1943. IEEE, 2017.

[4] Anthony Stentz. Optimal and efficient path planning for partially known environments. In *Intelligent Unmanned Ground Vehicles*, pages 203–220. Springer, 1997.

[5] Steven M LaValle. *Planning algorithms*. Cambridge University Press, 2006.

[6] Léonard Jaillet, Juan Cortés, and Thierry Siméon. Sampling-based path planning on configuration-space costmaps. *IEEE Transactions on Robotics*, 26(4):635–646, 2010.

[7] Ran Zhao. *Trajectory planning and control for robot manipulations*. PhD thesis, Université Paul Sabatier-Toulouse III, 2015.

[8] Mark W Spong, Seth Hutchinson, Mathukumalli Vidyasagar, et al. *Robot modeling and control*. John Wiley & Sons, Inc., 2006.

[9] Steven M LaValle. Rapidly-exploring random trees: A new tool for path planning. Technical report, Department of Computer Science, Iowa State University, 1998.

[10] Sertac Karaman and Emilio Frazzoli. Sampling-based algorithms for optimal motion planning. *The international journal of robotics research*, 30(7):846–894, 2011.

[11] Marios Xanthidis, Ioannis Rekleitis, and Jason M O'Kane. Rrt+: Fast planning for high-dimensional configuration spaces. *arXiv preprint arXiv:1612.07333*, 2016.

[12] Peter JM Van Laarhoven and Emile HL Aarts. Simulated annealing. In *Simulated annealing: Theory and applications*, pages 7–15. Springer, 1987.

[13] Dug Hun Hong and Chang-Hwan Choi. Multicriteria fuzzy decision-making problems based on vague set theory. *Fuzzy sets and systems*, 114(1):103–113, 2000.

[14] Robert Babuška and Henk Verbruggen. Neuro-fuzzy methods for nonlinear system identification. *Annual reviews in control*, 27(1):73–85, 2003.

[15] L-X Wang. Stable adaptive fuzzy control of nonlinear systems. *IEEE Transactions on fuzzy systems*, 1(2):146–155, 1993.

[16] L-X Wang. Fuzzy systems are universal approximators. In *[1992 Proceedings] IEEE International Conference on Fuzzy Systems*, pages 1163–1170. IEEE, 1992.

[17] L-X Wang and Jerry M Mendel. Generating fuzzy rules by learning from examples. *IEEE Transactions on systems, man, and cybernetics*, 22(6):1414–1427, 1992.

[18] L. . Wang. Training of fuzzy logic systems using nearest neighborhood clustering. In *[Proceedings 1993] Second IEEE International Conference on Fuzzy Systems*, pages 13–17 vol.1, March 1993.

[19] Lantos Béla. *Robotok irányítása (Robot Control)*. Akadémiai K., 2002.

[20] MG Mohanan and Ambuja Salgoankar. A survey of robotic motion planning in dynamic environments. *Robotics and Autonomous Systems*, 100:171–185, 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 388-401 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Racial Categorization Methods: A Survey

Krina B. Gabani*, Mayuri A. Mehta, Stephanie Noronha

Department of Computer Engineering, Sarvajanik College of Engineering and Technology, Surat, 395001, Gujarat, India

ARTICLE INFO

Article history:
Received: 30 January, 2020
Accepted: 25 May, 2020
Online: 11 June, 2020

Keywords:
Ethnicity
Ethnic group
Race
Racial classification
Anthropometry
Facial features
Race identification
Face recognition
Machine learning

ABSTRACT

Face explicitly provides the direct and quick way to evaluate human soft biometric information such as race, age and gender. Race is a group of human beings who differ from human beings of other races with respect to physical or social attributes. Race identification plays a significant role in applications such as criminal judgment and forensic art, human computer interface, and psychology science-based applications as it provides crucial information about the person. However, categorizing a person into respective race category is a challenging task because human faces comprise of complex and uncertain facial features. Several racial categorization methods are available in literature to identify race groups of humans. In this paper, we present a comprehensive and comparative review of these racial categorization methods. Our review covers survey of the important concepts, comparative analysis of single model as well as multi model racial categorization methods, applications, and challenges in racial categorization. Our review provides state-of-the-art technical information concerning racial categorization and hence, will be useful to the research community for development of efficient and robust racial categorization methods.

1. Introduction

Human face expresses social information that is highly useful in automated systems. It provides soft biometric information of human such as race, gender, age, identity and emotions [1-7]. This information is significant in interdisciplinary research areas such as psychology science, computer vision science, neuroscience, anthropological science as well as in the social security and forensic art department. Amongst the various types of biometric information, race information is crucial and is required for a wide range of applications. Race is a group of human beings differentiated based on physical or social attributes. Race conveys social and cultural traits of different communities. Facial features such as eyes, eyebrow, ear, nose, cheek, mouth, chin, forehead area and jaw differ from human to human and are highly dependent on racial category [8-14]. Figure 1 shows the difference in facial features of different racial groups.

Race analysis is essential in contemporary applications such as criminal judgment and forensic art [15-20], aesthetic surgery [21], healthcare [22-26], medico legal [27-29], video security surveillance and public safety [30], human computer interface [31-33] and face recognition [34]. In such applications, race analysis is required for identification of individuals. Several racial categorization methods have been proposed in literature. They are either single model racial categorization methods or multi model racial categorization methods. Single model racial categorization method uses facial features to recognize race [35-45]. Conversely, multi model racial categorization method considers fusion of physical characteristics such as gait pattern and audio clues in addition to facial features [5, 7, 46-50]. The majority of the practical applications involve single model racial categorization because facial data is available in large quantities compared to gait pattern and audio clues. However, there are applications that consider gait pattern and audio cues in absence of facial image. The single model racial categorization methods mainly differ from each other with respect to classification approach such as Support Vector Machine (SVM) [51-54], Convolutional Neural Network (CNN) [38, 44, 55-59], Artificial Neural Network (ANN) [60], Local Binary Pattern (LBP) [61] and Local Circular Pattern (LCP) [62]. In literature, participants based racial categorization methods such as diffusion model [63] and implicit racial attitude [64] are also available. The multi model racial categorization methods involve classification approaches such as SVM [65-67], logistic regression [66], Adaboost [66], random forest [66], CNN [68] and Haar-LBP histogram [69].

*Krina B. Gabani, Email: krinagabani007@gmail.com

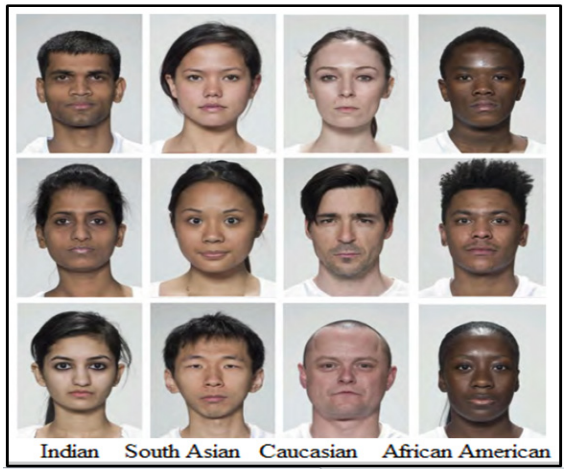


Figure 1: Illustrative facial characteristic variance of human races. (Figure source: http://www.faceresearch.org/)

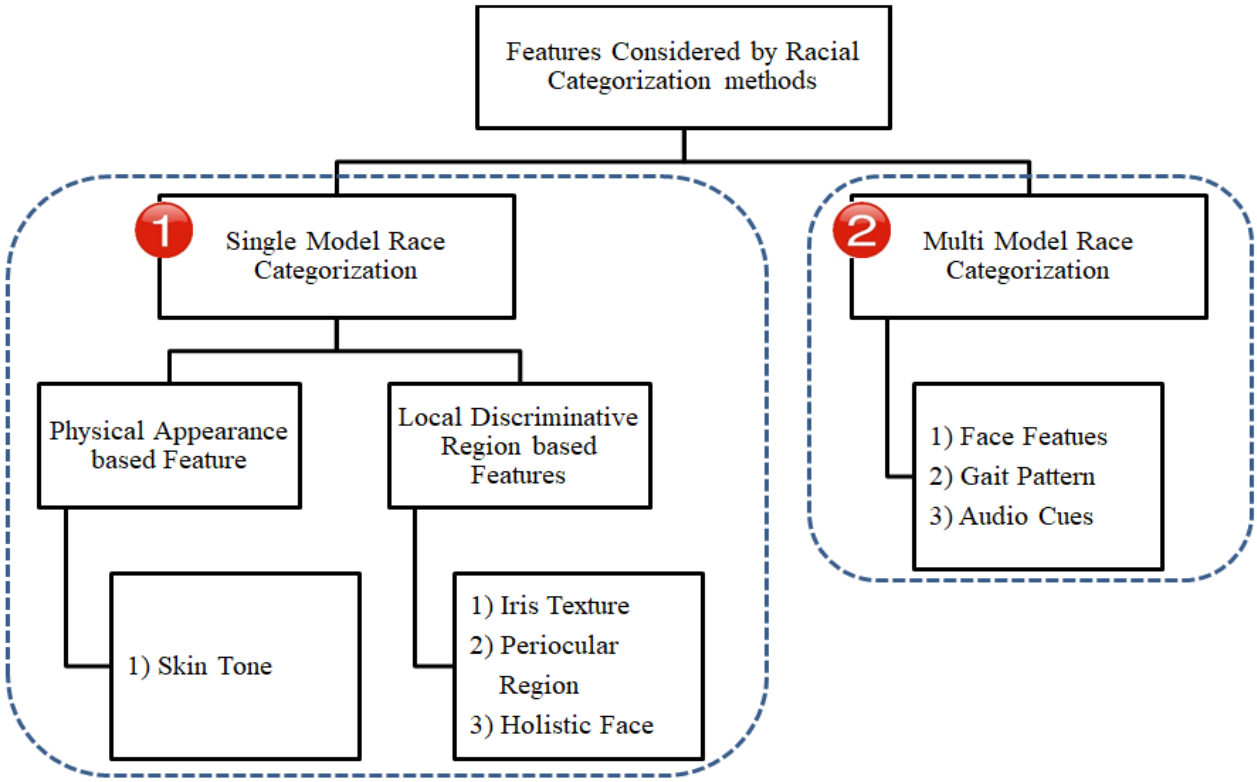
In this paper, we present a thorough and extensive study of various racial categorization methods. First, we present a taxonomy of available racial categorization methods. Then we describe several single model and multi model racial categorization methods. Based on our study, we identify several parameters to evaluate them. Subsequently, we present parametric evaluation of single model racial categorization methods and multi model racial categorization methods separately based on identified parameters. Next, we illustrate applications of racial categorization and future research direction in the field of racial categorization. Our comprehensive and comparative survey will serve as a catalogue to researchers in this area.

The rest of the paper is structured as follows: In section 2, we describe the taxonomy of racial categorization methods and features considered by different racial categorization methods. In section 3, we illustrate various single model racial categorization methods and their parametric evaluation. In section 4, we illustrate various multi model racial categorization methods and their parametric evaluation. Section 5 describes the major applications of racial categorization. In section 6, we list key challenges in the field of racial categorization. Finally, section 7 specifies conclusion and feature scope in the field of racial categorization.

2. Classification of Racial Categorization Methods

Racial categorization methods are broadly categorized into two categories: single model and multi model. As shown in Figure 2, the features used by single model and multi model methods to classify humans into features or local discriminative region based features or combination of both. Amongst the discriminative region based features, iris texture, periocular region or/and holistic face are used for racial categorization [70-77]. Multi model racial categorization takes into consideration face features, gait pattern and audio cues to classify humans into various race categories. Gait pattern is also useful to recognize the biometric information of humans [6, 78-80].

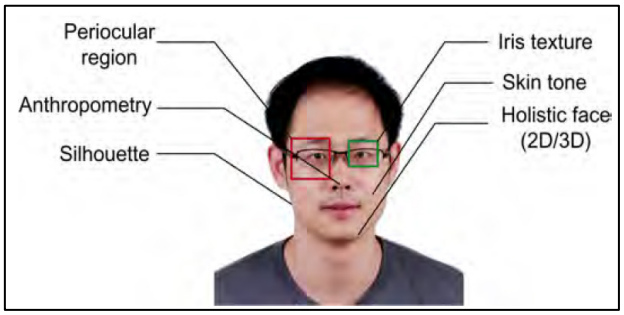
At this juncture, we clarify that at the top of all categorizations, a human is mainly divided into two categories, namely race and ethnicity, on the basis of his/her physical appearance and social appearance respectively. However, some researchers use words race and ethnicity interchangeably [24, 35, 81].



: Features considered for racial categorization

2.1. Features Considered by Single Model Racial Categorization

Figure 3 shows the facial features such as periocular region, anthropometry distance, silhouette, iris texture, skin tone and holistic face.



Skin tone differs mainly due to geographical location of humans. African-American, South-Asian, East-Asian, Caucasian, Indian and Arabian have different skin tones. Skin tone plays a minor role in identification of racial groups because skin color may also differ due to varying lighting conditions during the image capturing process [56, 64].

Like fingerprints, iris texture is a significant biometric characteristic of humans because it is unique for every human being [82]. It is highly useful for racial categorization because different race groups such as American, Indian and so on have different iris texture [59, 74-76, 82-84]. The key limitation of this feature is that it cannot be considered if race is to be identified from video because video may be of low quality and hence, may not give precise iris texture information [85-86].

Periocular region is defined as the region surrounded by eye. It is a region that overlays eyebrows, eyelid, eyelash and canthus [52]. It gives rich texture and biometric information as compared to iris texture [30]. Some facial features get influenced due to different facial poses and expressions. However, the periocular region does not get affected due to facial poses and expressions. Hence, it is considered as the most reliable feature for racial categorization [52, 62, 77].

Holistic face provides the texture information of various facial features such as eyes, nose, mouth, cheek, chin, skin color and jaw line [43, 51, 63, 87-89]. Extra frontal face features such as hairline and hair color in combination with cropped aligned face features ease racial categorization process [54].

2.2. Features Considered by Multi Model Racial Categorization

Multi model racial categorization improves accuracy of racial categorization via fusion of facial features with other human features such as gait pattern and audio cues [90] (Figure 4). Below we discuss the features considered by multi model racial categorization.

Gait pattern, also known as the walking pattern, is a prominent biometric feature that varies from human to human and is used for identification of a person [91-92]. Advanced racial categorization methods use gait pattern fused with facial features for overall effective racial categorization [92-97]. For videos in which humans at near to moderate distance have been captured, facial features are sufficient to identify the race. However, for videos in which humans at far distance have been captured, gait pattern is highly useful to identify the race of human because facial features of humans at far distance are not clearly visible. Thus, fusion of gait pattern with facial features improves overall racial categorization accuracy [65].

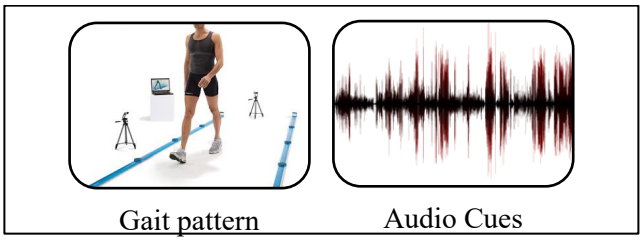


Figure 4: Features considered by multi model racial categorization

Audio pattern differs from race to race [98]. It is useful to identify race in case a video sequence or image of a person is not available. For instance, it is useful to identify race from phone calls.

3. Single Model Racial Categorization Methods

Race depends on physical and social characteristics of humans. As geographic distance increases, variation in facial features of inter-races become visible. As facial data is easily available compared to gait pattern, the majority of the racial categorization applications use a single model racial categorization method. Moreover, it has been revealed in literature that facial features are more prominent for race categorization [99-100]. Below we discuss various single model racial categorization methods available in literature.

3.1. Multi Ethnical Categorization using Manifold Learning

In [51], authors have proposed a method for intra-racial categorization based on facial landmarking. This method classifies eight intra-races residing in China based on facial landmarking concept. It includes Active Shape Model (ASM) to locate 77 facial landmarks. The landmarks are used to calculate three types of geometric facial features: distance, angle, and ratio. These features are provided to different classifiers such as Bayesian Net, Naive Bayesian, SMO, J4.8, RBF Network and LibSVM to identify the race category. The dimensionality reduction process carried out by manifold learning approach is useful to reduce the complexity. Though this method is efficient, it is not useful to identify race from a person's profile face images.

3.2. GWT and Ratina Sampling based Ethnicity Categorization

Multiclass SVM based ethnicity categorization method is proposed in [52]. Figure 5 shows the key steps involved in this method.

As shown in figure, first image is normalized via applying rotation operation and changing resolution. The resolution of the image is changed in such a manner that it maintains distance of 28 pixels between two inner corners of eyes. Subsequently, eye and mouth facial features are extracted by fusion of Gabor

Wavelet Transfer (GWT) and retina sampling for efficient categorization. GWT is used to extract accurate orientation and frequency of facial features. Retina sampling method is used to set facial feature points. The features are fed to multiclass SVM classifier for ethnicity identification. Typically eye is considered as a most prominent feature for racial categorization due to its pose invariant characteristic. On the other hand, uncertainty is introduced by mouth region due to its pose variant characteristic. The disadvantage of this method is that GWT provides erroneous features in case of hollow around the eyes. Moreover, Gabor features reflect error due to variation in frequency of eyelashes.

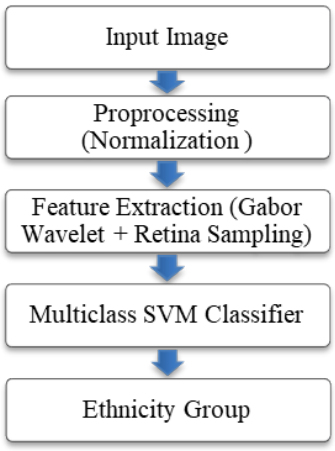


Figure 5: Steps for ethnicity categorization using GWT and retina sampling

3.3. Real Time Racial Categorization

A new method for racial categorization by fusion of Principal Component Analysis (PCA) and Independent Component Analysis (ICA) is introduced in [53]. It consists of major two steps: feature extraction and classification. During the feature extraction step, facial features are extracted using PCA. Subsequently, ICA is used to map and generate new facial features from facial features generated by PCA. New facial features are more suitable for efficient racial categorization. During the classification step, SVM classifier is applied in conjunction with '321' algorithm to classify races. '321'algorithm is inspired by the bootstrap approach for real time racial categorization from video streams. The categorization accuracy of this method can further be enhanced by including pre-processing step to diminish noise from the image and for face alignment.

3.4. Binary Tree based SVM for Ethnicity Detection

In this method, fusions of texture and shape facial features have been considered for better ethnicity categorization [54]. Figure 6 shows the functioning of this method. The first step pre-processing involves the operations such as image resize, image enhancement and image conversion. Then texture features are extracted using Gabor filter and shape features are extracted using Histogram Oriented Gradient (HOG). Subsequently, texture features and shape features are fused together. The fused feature vector is large and it requires more computational time. Hence, Kernel Principle Component Analysis (KPCA) algorithm is applied to reduce dimensionality and complexity. Fused facial features are given as input to binary tree based SVM for ethnicity detection.

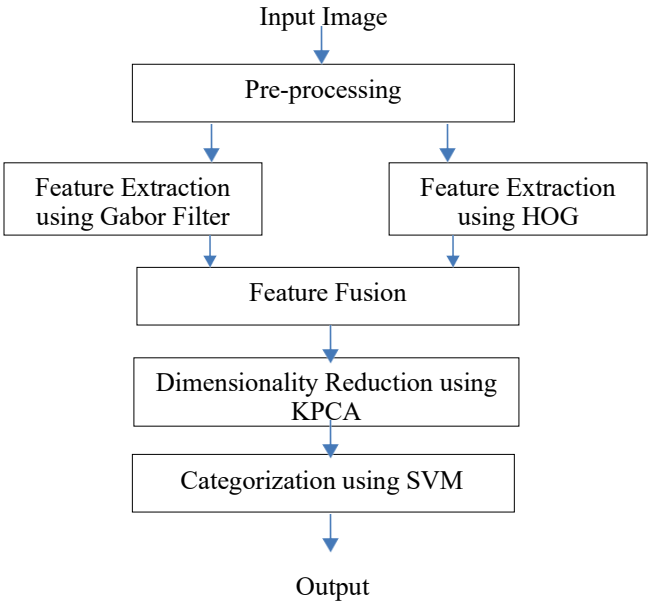


Figure 6: Steps for ethnicity detection using binary tree based SVM [54]

3.5. Racial Categorization using CNN

In [55], a hybrid supervised deep learning based racial categorization method has been proposed. It uses VGG 16 convolution neural network for facial feature extraction and categorization. 224 X 224 face image is given as an input to VGG-16 network for race prediction. Any CNN requires millions of images for training from scratch, which is critical a situation for the medical domain. Hence, to overcome an issue of small dataset, authors have used hybrid approach via fusing VGG 16 with image ranking engine to improve race prediction. It has been shown that image ranking engines work efficiently with CNN based classifiers even for small dataset. The fused feature information extracted by CNN and image ranking engine is used by SVM to learn racial class labels. This hybrid method provides better categorization accuracy.

Begin

1. Face detection using cascade classifier
2. Mark different facial features like nose, eyes and mouth.
3. Calculate distance and ratio between the marked facial features.
4. Detect different geometric facial features.
5. Detect skin color using YCbCrcolor model.
6. Detect forehead area using Sobel edge detection.
7. Normalize the forehead area considering coordinates of face and eyes and by applying following equation.

$$\text{Normalized Forehead Area} = \frac{\text{ForeheadArea}}{\text{TotalFaceArea}}$$

8. Create normalized feature matrix.
9. Train and validate neural network using feature matrix.
10. Test neural network for racial categorization.

End

Figure 7: Steps involved in racial categorization using ANN

3.6. Neural Network based Racial Categorization

In [56], skin color, forehead area, sobel edge and geometric features are fused for efficient race estimation. Authors have proposed two methods: 1) using Artificial Neural Network (ANN) and 2) using convolution neural network. The steps involved in racial categorization using ANN are shown in Figure 7.

CNN based racial categorization method uses pre-trained VGGNet for racial categorization. It has been observed by authors that CNN based method gives more accurate racial prediction for the given image as compared to ANN based method.

3.7. Local Circular Pattern for Race Identification

Local circular pattern for race identification method works on texture and shape features extracted from 2D face and 3D face respectively. A local circular pattern is an advanced version of a local binary pattern produced for feature extraction. LCP improves the widely utilized LBP and its variants by replacing binary quantization with clustering approach. As compared to LBP, LCP provides higher accuracy even for noisy data. Moreover, AdaBoost algorithm is used for selection of better features and thereby to improve the categorization accuracy. Experimental results have revealed that this method is time efficient and memory efficient.

3.8. Biometric based Machine Learning Method

In [62], authors have proposed a method that focuses on eye region features for racial categorization. The method comprises of five major steps shown in Figure 8. First facial coordinates are located using the DLib library. Subsequently, the region of interest (eye region) is extracted. Then features extracted by LBP and HOG are integrated for efficient racial categorization. LBP and HOG both are individually useful to extract features for categorization. However, fusion of LBP features with HOG features gives higher accuracy compared to other feature fusion approaches. The performance is tested utilizing different classifiers such as SVM, Multi-Layer Perceptron (MLP) and Quadratic Discriminant Analysis (QDA).

Begin

1. Locate facial coordinates using DLib library.
2. Extract Region of Interest (ROI) – eye region.
3. Extract features of eye region using LBP and HOG.
4. Fuse features extracted by LBP and HOG.
5. Input fused features to classifier for racial categorization.

End

Figure 8: Steps of biometric based machine learning method for racial categorization

3.9. Diffusion Model and Implicit Racial Attitude for Racial Group Identification

In [63-64, 101], manual racial categorization method is proposed. Race is identified by performing several tasks with participants. Diffusion model is used to identify response time boundaries of different participants. This method takes into consideration the visualization of participants for their own race and other races. Skin color is a less effective feature for automated racial categorization methods due to lightning conditions. However, it is a prominent feature for manual race prediction.

3.10. Performance Evaluation of Single Model Racial Categorization Methods

The above discussed race categorization methods are automated except the last one which is manual race categorization method. With increasing technology, manual race categorization is less effective and less useful as compared to automated racial categorization. Amongst the various automated racial categorization methods, CNN based methods produce more accurate results [55-56] as it considers deep facial features for racial categorization. We observed that the intra-race categorization is not much focused by the researchers in their study.

Based on our study on aforementioned single model racial categorization methods, we have identified the following parameters to compare them: dataset used, racial/ethnic class considered, region of interest, feature extraction operator/s and classifiers used. Dataset refers to the source of data. It is either available online in the form of a standard dataset or it is self-generated. HOIP Database, FERET (Facial Recognition Technology), FRGC v2.0 and BU-3DFE are examples of standard dataset. Self-generated dataset is created by researchers if their predefined requirements are not satisfied with a standard dataset. Racial/ethnic class represents the racial group targeted for study such as Indian, Chinese, Asian, European, African, African American, Caucasian, Bangladeshi, Mongolian, Caucasian, Negro, Hispanic and Pacific Islander. Region of interest specifies the area of face considered for racial categorization. Different face areas include eyes, eyebrow, ear, nose, cheek, mouth, chin, forehead area and jaw line. Some methods also consider skin color for race prediction. Feature extraction operator/s extracts the facial feature from the image. GWT, ICA, PCA, HOG, LBP, LCP are the features extraction operators used by different methods. Classifier specifies the classification algorithm used by the racial categorization method. Different classifiers such as SVM, CNN, ANN, kernel PCA, MLP, LDA, QDA and Kernel based Neural Network (KNN) have been used in different racial categorization methods. Table 1 presents the assessment of aforementioned single model racial categorization methods based these identified parameters.

4. Multi Model Racial Categorization Methods

Multi model racial categorization is highly useful when we do not have human's facial image information. It has been shown in literature that gait pattern and audio cues are amongst the prominent features for biometric information identification. Hence, multi model racial categorization methods use gait pattern, audio cues or fusion of facial features with gait pattern/audio cures to identify race. However, a smaller number of multi model racial categorization methods are available in literature because race data that includes gait pattern or audio cues is not available easily.

4.1. Multi-view Fused Gait based Ethnicity Classification

In [102], authors have proposed a method that identifies ethnicity from seven gait patterns captured from seven different

Table 1: Parametric Evaluation of Single Model Racial Categorization Methods

Method	Dataset used	Racial/ethnic class considered	Region of interest	Feature extraction Operator/s	Classifier/s used
Multi Ethnical Categorization using Manifold Learning [51]	Self-Generated	Chinese (8-Subgroups)	Face	-	Bayesian Net, Naive Bayesian, J4.8, RBF (Radial Basis Function) Network, LibSVM
GWT and Ratina Sampling based Ethnicity Categorization [52]	HOIP Database + Self- Generated	Asian, European, African	Eyes, Mouth	GWT, Ratina Sampling	SVM
Real Time Race Categorization [53]	FERET	Asian, Non-Asian	Face	ICA, PCA	SVM
Binary Tree based SVM for Ethnicity Detection [54]	FERET	Caucasian, African, Asian	Face	Gabor Filter, HOG	SVM, Kernel PCA
Racial Categorization using CNN [55]	Self- Generated	Bangladeshi, Chinese, Indian	Face	-	SVM
Neural Network based Racial Categorization [56]	FERET	Mongolian, Caucasian, Negro	Skin Colour, Forehead Area	-	ANN, CNN
Local Circular Pattern for Race Identification [61]	FRGC v2.0 and BU-3DFE	Whites and East-Asians	Face	LCP	Adaboost
Biometric based Machine Learning Method [62]	FERET	Asian, White, Black or African American, Hispanic, Pacific Islander, Native American	Eyes, Eyebrows, Periocular Region	LBP, HOG	SVM, MLP, LDA, QDA
Diffusion Model (Binary Classifier) [63]	Race Morph Sequence	Asian, Caucasian	Face	Manually	Manually
Implicit Racial Attitude [64]	Facial stimuli used in current research	African American, Caucasian	Skin Colour, Facial Physiognomy	Manually	Manually

angles. Figure 9 shows the key steps involved in the ethnicity identification process. First, all seven gait patterns are converted into corresponding Gait Energy Image (GEI). Next, seven GEIs are fused using three different fusion methods: score fusion, feature fusion and decision fusion. The goal of using three fusion methods is to accurately identify the ethnicity (race) of the person. Subsequently, features are extracted from the fused image using Multi-linear Principal Component Analysis (MPCA) and fed to classifier for ethnicity classification.

4.2. Hierarchical Fusion for Ethnicity Identification

In this method [65], gait pattern and facial features are fused for better ethnicity categorization. Figure 10 shows the two level processing involved in this method. First level involves gait pattern evolution. It takes gait video as an input. It includes the intermediate steps such as gait cycle estimation and GEI generation. First level also includes SVM for classification. Second level takes face video as an input. It comprises of three

Begin

1. Gait energy image generation of gait patterns captured from 7 different angles
2. Fusion of gait energy images
   - 2.1. Feature Level
   - 2.2. Score Level
   - 2.3. Decision Level
3. Feature extraction from fused image using MPCA
4. Ethnicity classification from extracted features

End

Figure 9: Steps involved in multi-view fused gait based ethnicity classification

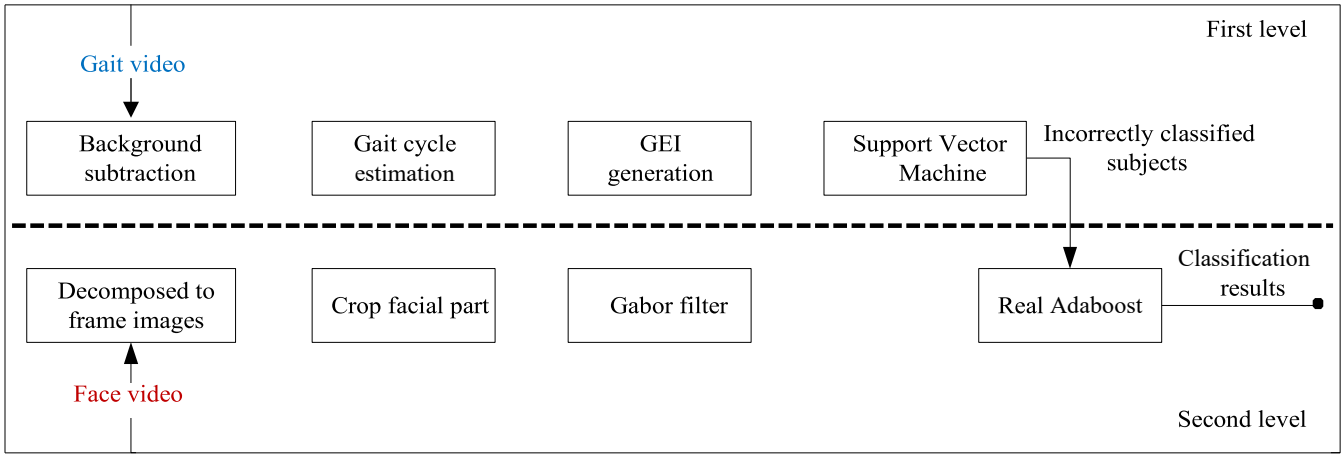


Figure 10: Steps of ethnicity categorization using hierarchical fusion system [65]

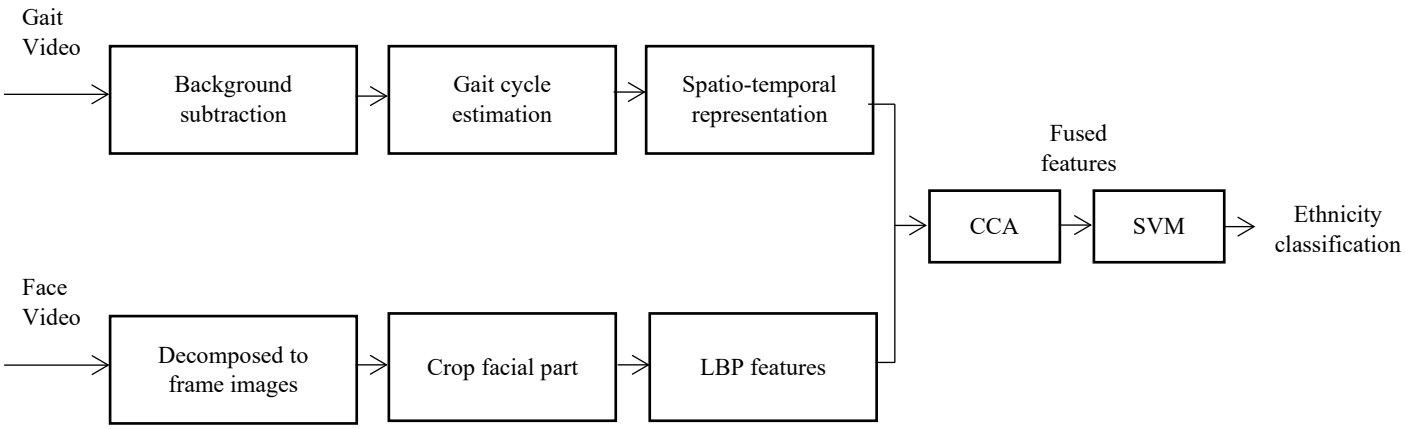


Figure 11: Block diagram of ethnicity classification system [67]

major steps: frame extraction from video, face detection and feature extraction using Gabor filter. Features extracted in first and second level are fused together to get accurate classification. The fused features are given to SVM and then to Adaboost to identify ethnicity.

4.3. Dialogue based Biometric Information Classification

In [66], a method for biometric information classification and deception detection has been proposed. It identifies gender, personality and ethnicity from the audio (dialogue). Lexical and acoustic-prosodic features are extracted from the dialogue. Lexical features are extracted using Linguistic Inquiry and Word Count (LIWC). Acoustic-prosodic features are extracted using Praat. Both types of features are given to different machine learning classifiers such as SVM, logistic regression, Adaboost and random forest for classification.

4.4. Cross-Model Biometric Matching

A new method for cross biometric matching by fusion of voice and facial image is introduced in [68]. The cross model is used for inferring the two types of information: 1) voice from human face and 2) human face from voice. This method involves two key steps: feature extraction and classification. The features are extracted from image as well as voice. The extracted features are fused and given as input to CNN for biometric matching and classification.

4.5. Gait and Face Fusion for Ethnicity Classification

Ethnicity classification system is proposed in [67]. It considers fusion of facial features and gait pattern for better classification. As shown in Figure 11, inputs to this system are gait video and facial video. Both videos are processed in parallel. During gait video processing, first background is subtracted from the video and subsequently each gait cycle pattern is estimated. Next all gait cycle patterns are represented using spatio- temporal representation for gait pattern characterization. During face video processing, frames are extracted from the face video and subsequently the facial part is cropped from the face image. Facial features are extracted from each frame using LBP. Features extracted from gait and face videos are fused together using Canonical Correlation Analysis (CCA). Fused features are given as input to SVM for ethnicity identification.

4.6. Performance Evaluation of Multi Model Racial Categorization Methods

As illustrated in Table 2, we identified the same set of parameters for comparison of multi model racial categorization methods as we identified for single model racial categorization methods. As defined and discussed previously in section 3, they are dataset used, biometric information considered for classification, region of interest, feature extraction operator/s and classifiers used in the method.

Table 2: Parametric Evaluation of Multi Model based Racial Categorization Methods

Method	Dataset used	Biometric information considered	Region of interest	Feature extraction operator/s	Classifier/s used
Multi-view Fused Gait based Ethnicity Classification [102]	Self- Generated	Ethnicity (East-Asian and South-American)	Gait Pattern	Multi-linear Principal Component Analysis	-
Hierarchical Fusion for Ethnicity Identification [65]	Self- Generated	Ethnicity (East-Asian and South-American)	Gait Pattern and Face	Gabor Filter	SVM
Dialogue based Classification [66]	NEO-FFI	Gender, Ethnicity and Personality	Dialogue (Audio)	LIWC and Praat	SVM, Logistic Regression, AdaBoost and Random Forest
Cross-Modal Biometric Matching [68]	VGGFace (Face data) and VoxCeleb (Audio Data)	Gender, Age, Ethnicity and Identity	Audio and Face	-	CNN
Gait and Face Fusion for Ethnicity Classification [67]	Self- Generated	Ethnicity	Gait Pattern and Face	LBP	SVM

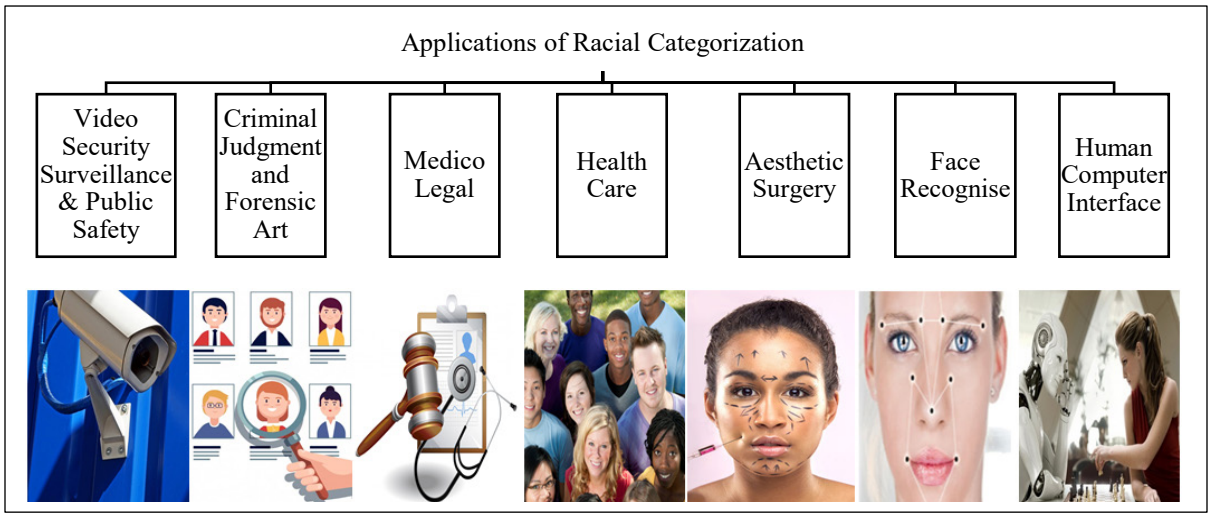


Figure 12: Application areas of racial categorization

5. Application of Racial Categorization

Racial categorization has a high impact on our social life. Race defines common physical characteristics of humans to represent his existence. Physical characteristics of humans of different races differ from each other. Racial categorization is significant for several applications. As shown in Figure 12, the major application areas are video security surveillance and public safety, criminal judgment and forensic art, medico legal, healthcare, aesthetic surgery, face recognition, and human computer interface.

5.1. Video Security Surveillance and Public Safety

Race identification from the subject's face plays a crucial role in video security surveillance. Video security surveillance system assists in identifying criminals by comparing the detected subject's image with the existing criminal database. Automated race identification system fused with video security surveillance system provides quick information about the subject [53]. Such a fused system is already in use at several airports and public places. Moreover, it has been proven useful for applications such as maritime, aviation, mass transformation, government office building, recreational centers, stadium and large retail malls.

5.2. Criminal Judgment and Forensic Art

Crime related investigation requires crucial information related to criminals including cross-country evidence (if any) [103-108]. Race/ethnicity of criminals provides such crucial information. Face is typically considered for criminal investigation because face conveys important information. Particularly, it conveys age, race and gender that are needed for criminal investigation. This information makes the investigation process easy for the government to find the right criminal [104,

109-111]. Moreover, such information is useful to prevent innocent people and provides justice to minority community groups.

Normally the forensic department has the subject's image captured using a public camera. However, it is difficult for the forensic department to manually extract the crucial information from the image. Conversely, racial categorization method can be used to identify the race from the image which assists in further investigation targeting a particular race community [106].

5.3. Medico Legal

Medico legal case is defined as a case of suffering or injury in which examination by the police is essential to determine the cause of suffering or injury. Suffering or injury may be due to several unnatural conditions such as accidents, burning and death. Race provides patient's information that is useful to law enforcing agency for further investigation [112]. By evaluating the race information of the medico legal case, law enforcing agency can obtain history of medico legal cases in that particular racial group [113-115]. Such information eases the investigation process and assists medico legal department in decision making.

5.4. Healthcare

Disease and healthcare issues are conflicting for different geographical areas due to their weather conditions, living sense and food. Healthcare treatment differs for different racial groups [116]. Thus, racial categorization is useful to solve the healthcare issues and to provide quick treatment [117-121]. Moreover, ethnic information is useful to provide appropriate services and special advantages to minority ethnic groups which are defined by the government for the minority and economically low conditions [122].

The center to Eliminate Health Disparities (CEHD) of the University of Texas Medical Branch (UTMB) has implemented the Information System for the health of people of UTMB to reduce disparities in health. Their information system is also known as REAL (race, ethnicity and language) [122]. Figure 13 illustrates the role of CEHD in the health system of UTMB and Galveston County as a whole. UTMB is a university health center that welcomes patients from diverse backgrounds. It provides services to different racial groups whose income level is below the poverty line. However, the main objectives of this REAL project are (1) to improve the UTMB's health information system for better diagnostics and stratified quality measures by race, ethnicity, language and status (2) to develop and disseminate contingency plans to address disparities through effective partnerships with relevant stakeholders.

5.5. Aesthetic Surgery

Aesthetic surgery is described as a facial plastic surgery either for the beautification of face or to create an attractive face [123]. Anthropometric measurement is the distance between two facial points. It has been revealed in literature that anthropological measurements such as ratio, geometric distance and Euclidean distance are different for different racial groups [124-126]. Geometric and Euclidean distances are the distances between primary or secondary facial landmarks on the frontal face/profile face. Depending on the race of patient, anthropological measurements are derived and used in aesthetic surgery [127-129]. For example, aesthetic surgery for Chinese people and Indian people is different as both groups have different facial features and thereby different anthropological measurements.

5.6. Face Recognition

Racial difference in humans is useful for biometric illustration and human identification [130]. Race cues and race wise anthropological measurements make it easy for face recognition systems to recognize the person [131-134]. Moreover, integration of race information with face recognition makes the face recognition system more intelligent and quick for accurate face recognition [135-136].

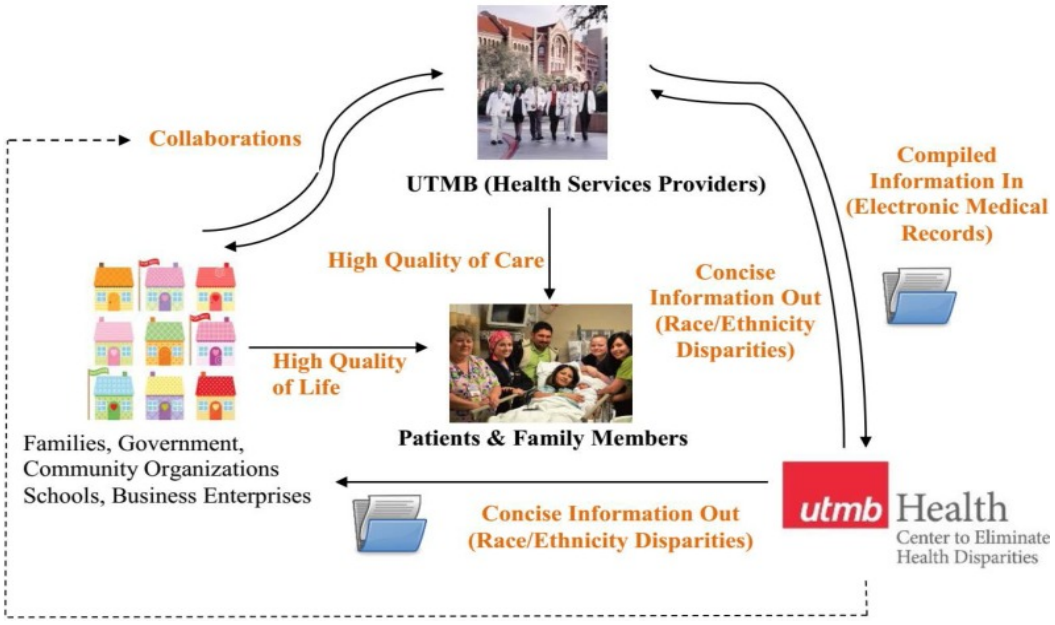


Figure 13: Role of the health center system to eliminate health disparities for different ethnicity [122]

5.7. Human Computer Interface

Nowadays, several systems are automated using robots [137-140]. Consumers of such robotic systems need to interact with robots frequently. In human-robot communication, racial cues play an important role [141-145]. Specifically, by recognizing the race of a human from his face, behavior and expression, robots can deliver the relevant services to humans. Such robotic systems are useful for an open service atmosphere where robots work as humans [146]. In particular, they are useful in hospitals, malls, stadiums, hotels, gaming zones and intelligent HCI organizations for easy communication with humans.

6. Challenges in Racial Categorization

Several challenges are faced to get correct and accurate racial categorization. Below we mention the major challenges faced in racial categorization. These challenges create new opportunities for researchers in this field to carry out further research. An efficient racial categorization method can be developed by overcoming these challenges and higher classification accuracy can be achieved.

6.1. Intra-race Categorization

To the best of our knowledge, intra-race categorization has not been focused much in literature. 85% of the worldwide population is divided into major 7-racial groups, namely African-American, South-Asian, East-Asian, Caucasian, Indian, Arabian and Latino race [30]. Intra-race categorization for aforementioned racial groups is challenging due to severe similarity in facial features and in physical appearance of humans belonging to a particular group [147-148]. It is difficult to infer different clues for intra-race categorization.

6.2. Anthropometry Measurements

Facial landmarking technique is used to measure anthropometry measurements. It has been shown in literature that accuracy of anthropometry measurements and thereby accuracy of racial categorization method varies with respect to the number of facial landmarks [21, 51, 56, 62, 147-148]. Thus, existing landmarking methods can be further improved by increasing the number of landmarks. Moreover, landmarks on forehead area, hairline and earlobe can be additionally considered to increase accuracy further [56]. In addition, anatomists have revealed that Ear pinna and Iannarelli's measures differ for different racial groups. Like finger print, ear pinna is unique for each individual.

6.3. Real-Time Data

It is required to process real-time video streams at public places such as airports, hospitals, health care centers, malls and stadiums for public safety and security systems. However, existing racial categorization methods are not applicable and reliable for processing real-time video stream [53]. They are applicable to only off-line image dataset.

6.4. Manual Racial Categorization

The aforementioned issues are related to automated racial categorization methods. However, the issues faced by manual racial categorization methods are different. The major issue related to manual racial categorization which involves participants is that the number of stimulus levels for race prediction is limited. Stimulus level is defined as the number of tasks performed by participants for race prediction [56, 61].

7. Conclusion

In this paper, we have presented in-depth review on various single model and multi model racial categorization methods. Moreover, parametric evaluation of racial categorization methods based on identified set of parameters is presented. It has been observed that fusion of facial features and physical appearance provides accurate race categorization. Moreover, it has also been observed that CNN based racial categorization model gives substantially higher accuracy because it extracts deep features. Our rigorous review on racial categorization methods will provide researchers state-of-the-art advancements related to racial categorization methods. Furthermore, the applications and challenges of racial categorization discussed herein will help researchers to develop an efficient and competent racial categorization method.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] S. Gutta, H. Wechsler and P. Phillips, "Gender and ethnic classification of face images", Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition. https://doi.org/10.1109/afgr.1998.670948.

[2] B. Xia, B. Ben Amor and M. Daoudi, "Joint gender, ethnicity and age estimation from 3D faces", Image and Vision Computing, vol. 64, pp. 90-102, 2017. https://doi.org/10.1016/j.imavis.2017.06.004.

[3] S. Azzakhnini, L. Ballihi and D. Aboutajdine, "Combining Facial Parts For Learning Gender, Ethnicity, and Emotional State Based on RGB-D Information", ACM Transactions on Multimedia Computing, Communications, and Applications, vol. 14, no. 1, pp. 1-14, 2018. https://doi.org/10.1145/3152125.

[4] P. Mukudi and P. Hills, "The combined influence of the own-age, -gender, and -ethnicity biases on face recognition", Acta Psychologica, vol. 194, pp. 1-6, 2019. https://doi.org/10.1016/j.actpsy.2019.01.009.

[5] Xiaoli Zhou and B. Bhanu, "Integrating Face and Gait for Human Recognition", 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06). https://doi.org/10.1109/CVPRW.2006.103.

[6] J. Yoo, D. Hwang and M. Nixon, "Gender Classification in Human Gait Using Support Vector Machine", Advanced Concepts for Intelligent Vision Systems, pp. 138-145, 2005. https://doi.org/10.1007/11558484_18.

[7] Ju Han and B. Bhanu, "Statistical feature fusion for gait-based human recognition", Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. https://doi.org/10.1109/CVPR.2004.1315252.

[8] O. Çeliktutan, S. Ulukaya and B. Sankur, "A comparative study of face landmarking methods," EURASIP Journal on Image and Video Processing, vol. 2013, no. 1, 2013. https://doi.org/10.1186/1687-5281-2013-13.

[9] Y. Wu and Q. Ji, "Facial Landmark Detection: A Literature Survey," International Journal of Computer Vision, 2018. https://doi.org/10.1007/s11263-018-1097-z.

[10] H. Akakin and B. Sankur, "Analysis of Head and Facial Gestures Using Facial Landmark Trajectories," Biometric ID Management and Multimodal Communication, pp. 105-113, 2009. https://doi.org/10.1007/978-3-642-04391-8_14.

[11] L. Farkas, M. Katic and C. Forrest, "International Anthropometric Study of Facial Morphology in Various Ethnic Groups/Races," Journal of Craniofacial Surgery, vol. 16, no. 4, pp. 615-646, 2005. https://doi.org/10.1097/01.scs.0000171847.58031.9e.

[12] H. Dibeklioglu, A. Salah and T. Gevers, "A Statistical Method for 2-D Facial Landmarking," IEEE Transactions on Image Processing, vol. 21, no. 2, pp. 844-858, 2012. https://doi.org/10.1109/TIP.2011.2163162.

[13] B. Martinez, M. Valstar, X. Binefa and M. Pantic, "Local Evidence Aggregation for Regression-Based Facial Point Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 5, pp. 1149-1163, 2013. https://doi.org/10.1109/TPAMI.2012.205.

[14] M. Taner Eskil and K. Benli, "Facial expression recognition based on anatomy," Computer Vision and Image Understanding, vol. 119, pp. 1-14, 2014. https://doi.org/10.1016/j.cviu.2013.11.002.

[15] E. Lloyd, K. Hugenberg, A. McConnell, J. Kunstman and J. Deska, "Black and White Lies: Race-Based Biases in Deception Judgments", Psychological Science, vol. 28, no. 8, pp. 1125-1136, 2017. https://doi.org/10.1177/0956797617705399.

[16] G. Breetzke, "The concentration of urban crime in space by race: evidence from South Africa", Urban Geography, vol. 39, no. 8, pp. 1195-1220, 2018. https://doi.org/10.1080/02723638.2018.1440127.

[17] W. Threadcraft-Walker and H. Henderson, "Reflections on race, personality, and crime", Journal of Criminal Justice, vol. 59, pp. 38-41, 2018. https://doi.org/10.1016/j.jcrimjus.2018.05.005.

[18] C. Lehman, T. Hammond and G. Agyemang, "Accounting for crime in the US: Race, class and the spectacle of fear", Critical Perspectives on Accounting, vol. 56, pp. 63-75, 2018. https://doi.org/10.1016/j.cpa.2018.01.002.

[19] B. Jefferson, "Predictable Policing: Predictive Crime Mapping and Geographies of Policing and Race", Annals of the American Association of Geographers, vol. 108, no. 1, pp. 1-16, 2017. https://doi.org/10.1080/24694452.2017.1293500.

[20] C. Petsko and G. Bodenhausen, "Race–Crime Congruency Effects Revisited: Do We Take Defendants' Sexual Orientation into Account?," Social Psychological and Personality Science, vol. 10, no. 1, pp. 73-81, 2017. https://doi.org/10.1177/1948550617736111.

[21] M. Wai et al., "Nasofacial Anthropometric Study among University Students of Three Races in Malaysia", 2019. https://doi.org/10.1155/2015/780756.

[22] E. Rata and C. Zubaran, "Ethnic Classification in the New Zealand Health Care System", Journal of Medicine and Philosophy, vol. 41, no. 2, pp. 192-209, 2016. https://doi.org/10.1093/jmp/jhv065.

[23] J. Reid, D. Cormack and M. Crowe, "The significance of socially-assigned ethnicity for self-identified Māori accessing and engaging with primary healthcare in New Zealand", Health: An Interdisciplinary Journal for the Social Study of Health, Illness and Medicine, vol. 20, no. 2, pp. 143-160, 2015. https://doi.org/10.1177/1363459315568918.

[24] S. Mehta et al., "Race/Ethnicity and Adoption of a Population Health Management Approach to Colorectal Cancer Screening in a Community-Based Healthcare System", Journal of General Internal Medicine, vol. 31, no. 11, pp. 1323-1330, 2016. https://doi.org/10.1007/s11606-016-3792-1

[25] .L. Culley, "Transcending transculturalism? Race, ethnicity and health-care", Nursing Inquiry, vol. 13, no. 2, pp. 144-153, 2006. https://doi.org/10.1111/j.1440-1800.2006.00311.x.

[26] M. Gibbons, " Use of Health Information Technology among Racial and Ethnic Underserved Communities," Perspect Health Inf Manag, vol. 8, 2011.

[27] A. Memarian, S. Mostafavi, E. Zarei, S. Esfahani and B. Ghorbanzadeh, "The Credibility of Cephalogram Parameters in Gender Identification From Medico-Legal Relevance Among the Iranian Populatio," International Journal of Medical Toxicology and Forensic Medicine, vil. 9, 2019.

[28] M. Ibrahim, A. Khalifa, H. Hassan, H. Tamam and A. Hagras, "Estimation of stature from hand dimensions in North Saudi population, medicolegal view," The Saudi Journal of Forensic Medicine and Sciences, vol. 1, pp. 19-27, 2018. https://doi.org/10.4103/sjfms.sjfms_10_17.

[29] B. Bukovitz, J. Meiman, H. Anderson and E. Brooks, "Silicosis: Diagnosis and Medicolegal Implications", Journal of Forensic Sciences, 2019. https://doi.org/10.1111/1556-4029.14048.

[30] S. Fu, H. He and Z. Hou, "Learning Race from Face: A Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 12, pp. 2483-2509, 2014. . https://doi.org/10.1109/TPAMI.2014.2321570.

[31] J. Carroll, "Evaluation, Description and Invention: Paradigms for Human-Computer Interaction", Advances in Computers, pp. 47-77, 1989. https://doi.org/10.1016/S0065-2458(08)60532-X.

[32] M. Mackal, "The effects of ethnicity and message content on affective outcomes in a computer-based learning environment," 2009.

[33] V. Evers, "Human - Computer Interfaces: Designing for Culture," 1997.

[34] B. Marsh, K. Pezdek and D. Ozery, "The cross-race effect in face recognition memory by bicultural individuals", Acta Psychologica, vol. 169, pp. 38-44, 2016. https://doi.org/10.1016/j.actpsy.2016.05.003.

[35] L. Farkas, M. Katic and C. Forrest, "International Anthropometric Study of Facial Morphology in Various Ethnic Groups/Races", Journal of Craniofacial Surgery, vol. 16, no. 4, pp. 615-646, 2005. https://doi.org/10.1097/01.scs.0000171847.58031.9e.

[36] S. Ozdemir, D. Sigirli, I. Ercan and N. Cankur, "Photographic Facial Soft Tissue Analysis of Healthy Turkish Young Adults: Anthropometric Measurements", Aesthetic Plastic Surgery, vol. 33, no. 2, pp. 175-184, 2008. https://doi.org/10.1007/s00266-008-9274-z.

[37] E. Boutellaa, A. Hadid, M. Bengherabi and S. Ait-Aoudia, "On the use of Kinect depth data for identity, gender and ethnicity classification from facial images", Pattern Recognition Letters, vol. 68, pp. 270-277, 2015. https://doi.org/10.1016/j.patrec.2015.06.027.

[38] N. Narang and T. Bourlai, "Gender and ethnicity classification using deep learning in heterogeneous face recognition", 2016 International Conference on Biometrics (ICB), 2016. https://doi.org/10.1109/ICB.2016.7550082.

[39] Y. Wang, X. Duan, X. Liu, C. Wang and Z. Li, "Semantic description method for face features of larger Chinese ethnic groups based on improved WM method", Neurocomputing, vol. 175, pp. 515-528, 2016. https://doi.org/10.1016/j.neucom.2015.10.089.

[40] L. Zhao and S. Bentin, "Own- and other-race categorization of faces by race, gender, and age", Psychonomic Bulletin & Review, vol. 15, no. 6, pp. 1093-1099, 2008. https://doi.org/10.3758/PBR.15.6.1093.

[41] X. Lu, H. Chen and A. Jain, "Multimodal Facial Gender and Ethnicity Identification", Advances in Biometrics, pp. 554-561, 2005. https://doi.org/10.1007/11608288_74.

[42] G. Guo and G. Mu, "A framework for joint estimation of age, gender and ethnicity on a large database", Image and Vision Computing, vol. 32, no. 10, pp. 761-770, 2014. https://doi.org/10.1016/j.imavis.2014.04.011.

[43] A. da Silva, L. Magri, L. Andrade and M. da Silva, "Three-dimensional analysis of facial morphology in Brazilian population with Caucasian, Asian, and Black ethnicity", Journal of Oral Research and Review, vol. 9, no. 1, p. 1, 2017. http://dx.doi.org/10.1590/0103-6440201802027.

[44] T. Vo, T. Nguyen and C. Le, "Race Recognition Using Deep Convolutional Neural Networks", Symmetry, vol. 10, no. 11, p. 564, 2018. https://doi.org/10.3390/sym10110564.

[45] D. Bobeldyk and A. Ross, "Analyzing Covariate Influence on Gender and Race Prediction From Near-Infrared Ocular Images", IEEE Access, vol. 7, pp. 7905-7919, 2019. https://doi.org/10.1109/ACCESS.2018.2886275.

[46] S. Samangooei and M. Nixon, "Performing content-based retrieval of humans using gait biometrics", Multimedia Tools and Applications, vol. 49, no. 1, pp. 195-212, 2009. https://doi.org/10.1007/s11042-009-0391-8.

[47] A. Kale, A. Roychowdhury and R. Chellappa, "Fusion of gait and face for human identification", 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing. https://doi.org/10.1109/ICASSP.2004.1327257.

[48] S. Sarkar, P. Phillips, Z. Liu, I. Vega, P. Grother and K. Bowyer, "The humanID gait challenge problem: data sets, performance, and analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 2, pp. 162-177, 2005. https://doi.org/10.1109/TPAMI.2005.39.

[49] Ju Han and Bir Bhanu, "Individual recognition using gait energy image", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 2, pp. 316-322, 2006. https://doi.org/10.1109/TPAMI.2006.38.

[50] J. Gomez-Ezeiza, J. Torres-Unda, N. Tam, J. Irazusta, C. Granados and J. Santos-Concejero, "Race walking gait and its influence on race walking economy in world-class race walkers", Journal of Sports Sciences, vol. 36, no. 19, pp. 2235-2241, 2018. https://doi.org/10.1080/02640414.2018.1449086.

[51] C. Wang, Q. Zhang, X. Duanand J. Gan, "Multi-ethnical Chinese facial characterization and analysis," Multimedia Tools and Applications,2018. https://doi.org/10.1007/s11042-018-6018-1.

[52] S. Hosoi, E. Takikawa and M. Kawade, "Ethnicity estimation with facial images," Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. https://doi.org/10.1109/AFGR.2004.1301530.

[53] O. Yongsheng, W. Xinyu, Q. Huihuanand Yangsheng Xu, "A real time race categorization system," IEEE International Conference on Information Acquisition, 2005.

[54] M. Uddin and S. Chowdhury, "An integrated approach to classify gender and ethnicity," 2016 International Conference on Innovations in Science, Engineering and Technology (ICISET), 2016. https://doi.org/10.1109/ICISET.2016.7856480.

[55] Z. Heng, M. Dipu and K. Yap, "Hybrid Supervised Deep Learning for Ethnicity Categorization using Face Images," 2018 IEEE International Symposium on Circuits and Systems (ISCAS), 2018.

[56] S. Masood, S. Gupta, A. Wajid, S. Gupta and M. Ahmed, "Prediction of Human Ethnicity from Facial Images Using Neural Networks," Advances in Intelligent Systems and Computing, pp. 217-226, 2017. https://doi.org/10.1007/978-981-10-3223-3_20.

[57] A. Das, A. Dantcheva and F. Bremond, "Mitigating Bias in Gender, Age and Ethnicity Classification: a Multi-Task Convolution Neural Network Approach," ECCV, 2018. https://link.springer.com/conference/eccv.

[58] N. Srinivas, H. Atwal, D. Rose, G. Mahalingam, K. Ricanek and D. Bolme, "Age, Gender, and Fine-Grained Ethnicity Prediction Using Convolutional Neural Networks for the East Asian Face Dataset", 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), 2017. https://doi.org/10.1109/FG.2017.118.

[59] A. Mohammad and J. Al-Ani, "Convolutional Neural Network for Ethnicity Classification using Ocular Region in Mobile Environment", 2018 10th Computer Science and Electronic Engineering (CEEC), 2018. https://doi.org/10.1109/CEEC.2018.8674194.

[60] C. Bagchi, D. Geraldine Bessie Amali and M. Dinakaran, "Accurate Facial Ethnicity Classification Using Artificial Neural Networks Trained with Galactic Swarm Optimization Algorithm", Advances in Intelligent Systems and Computing, pp. 123-132, 2018. https://doi.org/10.1007/978-981-13-3329-3_12.

[61] D. Huang, H. Ding, C. Wang, Y. Wang, G. Zhang and L. Chen, "Local circular patterns for multi-modal facial gender and ethnicity categorization," 2018. https://doi.org/10.1016/j.imavis.2014.06.009.

[62] A. Mohammad and J. Al-Ani, "Towards ethnicity detection using learning based classifiers," 2017 9th Computer Science and Electronic Engineering (CEEC), 2017. https://doi.org/10.1109/CEEC.2017.8101628.

[63] C. Benton and A. Skinner, "Deciding on race: A diffusion model analysis of race-categorisation," 2018. https://doi.org/10.1016/j.cognition.2015.02.011.

[64] E. Stepanova and M. Strube, "The role of skin color and facial physiognomy in racial categorization: Moderation by implicit racial attitudes," Journal of Experimental Social Psychology, vol. 48, no. 4, pp. 867-878, 2012. https://doi.org/10.1016/j.jesp.2012.02.019.

[65] D. Zhang, Y. Wang and Z. Zhang, "Ethnicity Classification Based on a Hierarchical Fusion", Biometric Recognition, pp. 300-307, 2012. https://doi.org/10.1007/978-3-642-35136-5_36.

[66] S. Levitan, Y. Levitan, G. An, M. Levine, R. Levitan, A. Rosenberg and J. Hirschberg, "Identifying Individual Differences in Gender, Ethnicity, and Personality from Dialogue for Deception Detection," NAACL-HLT, pp. 40-44, 2016.

[67] D. Zhang, Y. Wang, Z. Zhang and M. Hu, "Ethnicity classification based on fusion of face and gait", 2012 5th IAPR International Conference on Biometrics (ICB), 2012. https://doi.org/10.1109/ICB.2012.6199781.

[68] A. Nagrani, S. Albanie and A. Zisserman, "Seeing Voices and Hearing Faces: Cross-Modal Biometric Matching," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[69] Hengliang Tang, Yanfeng Sun, Baocai Yin and Yun Ge, "Face recognition based on Haar LBP histogram", 2010 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE), 2010. https://doi.org/10.1109/ICACTE.2010.5579370.

[70] R. Sáenz and M. Morales, "6 Demography of Race and Ethnicity", Handbooks of Sociology and Social Research, pp. 163-207, 2019. https://doi.org/10.1007/978-3-030-10910-3_7.

[71] H. Qiu, S. Fang and K. Song, "Method and apparatus for face classification," Google pattern, 2018.

[72] C. Lv, Z. Wu, D. Zhang, X. Wang and M. Zhou, "3D Nose shape net for human gender and ethnicity classification", Pattern Recognition Letters, 2018. https://doi.org/10.1016/j.patrec.2018.11.010.

[73] M. Singh, S. Nagpal, M. Vatsa, R. Singh, A. Noore and A. Majumdar, "Gender and ethnicity classification of Iris images using deep class-encoder", 2017 IEEE International Joint Conference on Biometrics (IJCB), 2017.

[74] H. Jang, J. Yoon, Y. Kim and Y. Park, "Classification of Iris Colors and Patterns in Koreans", Healthcare Informatics Research, vol. 24, no. 3, p. 227, 2018. https://doi.org/10.4258/hir.2018.24.3.227.

[75] G. Mabuza-Hocquet, F. Nelwamondo and T. Marwala, "Ethnicity Distinctiveness Through Iris Texture Features Using Gabor Filters", Intelligent Information and Database Systems, pp. 551-560, 2017. https://doi.org/10.1007/978-3-319-54430-4_53.

[76] D. Bobeldyk and A. Ross, "Analyzing Covariate Influence on Gender and Race Prediction From Near-Infrared Ocular Images", IEEE Access, vol. 7, pp. 7905-7919, 2019. https://doi.org/10.1109/ACCESS.2018.2886275.

[77] H. Chen, M. Gao, K. Ricanek, W. Xu and B. Fang, "A Novel Race Classification Method Based on Periocular Features Fusion", International Journal of Pattern Recognition and Artificial Intelligence, vol. 31, no. 08, p. 1750026, 2017. https://doi.org/10.1142/S0218001417500264.

[78] S. Arseev, A. Konushin and V. Liutov, "Human Recognition by Appearance and Gait", Programming and Computer Software, vol. 44, no. 4, pp. 258-265, 2018. https://doi.org/10.1134/S0361768818040035.

[79] W. Chi, J. Wang and M. Meng, "A Gait Recognition Method for Human Following in Service Robots", IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 48, no. 9, pp. 1429-1440, 2018. https://doi.org/10.1109/TSMC.2017.2660547.

[80] I. Rida, A. Bouridane, G. Marcialis and P. Tuveri, "Improved Human Gait Recognition", Image Analysis and Processing ICIAP 2015, pp. 119-129, 2015. https://doi.org/10.1007/978-3-319-23234-8_12.

[81] P. Shah and N. Davis, "Comparing Three Methods of Measuring Race/Ethnicity", The Journal of Race, Ethnicity, and Politics, vol. 2, no. 1, pp. 124-139, 2017. https://doi.org/10.1017/rep.2016.27.

[82] R. Lee, B. Chon, S. Lin, M. He and S. Lin, "Association of Ocular Conditions With Narrow Angles in Different Ethnicities", American Journal of Ophthalmology, vol. 160, no. 3, pp. 506-515.e1, 2015. https://doi.org/10.1016/j.ajo.2015.06.002.

[83] B. Latinwo, A. Falohun, E. Omidiora and B. Makinde, "Iris Texture Analysis for Ethnicity Classification Using Self-Organizing Feature Maps", Journal of Advances in Mathematics and Computer Science, vol. 25, no. 6, pp. 1-10, 2018. https://doi.org/10.9734/JAMCS/2017/29634.

[84] G. Mabuza-Hocquet, F. Nelwamondo and T. Marwala, "Ethnicity Prediction and Classification from Iris Texture Patterns: A Survey on Recent Advances", 2016 International Conference on Computational Science and Computational Intelligence (CSCI), 2016. https://doi.org/10.1109/CSCI.2016.0159.

[85] M. Singh, S. Nagpal, M. Vatsa, R. Singh, A. Noore and A. Majumdar, "Gender and ethnicity categorization of Iris images using deep class-encoder," 2017 IEEE International Joint Conference on Biometrics (IJCB), 2017.

[86] S. Lagree and K. Bowyer, "Predicting ethnicity and gender from iris texture," 2011 IEEE International Conference on Technologies for Homeland Security (HST), 2011. https://doi.org/10.1109/THS.2011.6107909.

[87] K. Huri, E. (Omid) David and N. Netanyahu, "DeepEthnic: Multi-label Ethnic Classification from Face Images", Artificial Neural Networks and Machine Learning – ICANN 2018, pp. 604-612, 2018. https://doi.org/10.1007/978-3-030-01424-7_59.

[88] C. Wang, Q. Zhang, W. Liu, Y. Liu and L. Miao, "Facial feature discovery for ethnicity recognition", Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 9, no. 1, p. e1278, 2018. https://doi.org/10.1002/widm.1278.

[89] M. Sajid et al., "Facial Asymmetry-Based Anthropometric Differences between Gender and Ethnicity", Symmetry, vol. 10, no. 7, p. 232, 2018. https://doi.org/10.3390/sym10070232.

[90] G. Shakhnarovich, L. Lee and T. Darrell, "Integrated face and gait recognition from multiple views", Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001. https://doi.org/10.1109/CVPR.2001.990508.

[91] D. Zhang, Y. Wang and Z. Zhang, "Ethnicity Categorization Based on a Hierarchical Fusion," Biometric Recognition, pp. 300-307, 2012. https://doi.org/10.1007/978-3-642-35136-5_36.

[92] A. Mecca, "Impact of Gait Stabilization: A Study on How to Exploit it for User Recognition", 2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), 2018. https://doi.org/10.1109/SITIS.2018.00090.

[93] Y. Liu, K. Lu, S. Yan, M. Sun, D. Lester and K. Zhang, "Gait phase varies over velocities", Gait & Posture, vol. 39, no. 2, pp. 756-760, 2014. https://doi.org/10.1016/j.gaitpost.2013.10.009.

[94] Y. Wang, S. Yu, Y. Wang and T. Tan, "Gait Recognition Based on Fusion of Multi-view Gait Sequences", Advances in Biometrics, pp. 605-611, 2005. https://doi.org/10.1007/11608288_80.

[95] Xiaxi Huang and N. Boulgouris, "Gait Recognition using Multiple Views", 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, 2008. https://doi.org/10.1109/ICASSP.2008.4517957.

[96] A. Ross and A. Jain, "Information fusion in biometrics", Pattern Recognition Letters, vol. 24, no. 13, pp. 2115-2125, 2003. https://doi.org/10.1016/S0167-8655(03)00079-5.

[97] M. Deng and C. Wang, "Human gait recognition based on deterministic learning and data stream of Microsoft Kinect", IEEE Transactions on Circuits and Systems for Video Technology, pp. 1-1, 2018. https://doi.org/10.1109/TCSVT.2018.2883449.
[98] L. Sturm, K. Donahue, M. Kasting, A. Kulkarni, N. Brewer and G. Zimet, "Pediatrician-Parent Conversations About Human Papillomavirus Vaccination: An Analysis of Audio Recordings," Journal of Adolescent Health, vol. 61, no. 2, pp. 246-251, 2017. https://doi.org/10.1016/j.jadohealth.2017.02.006.
[99] A. Rehman, G. Khan, A. Siddiqi, A. Khan and U. Khan, "Modified Texture Features from Histogram and Gray Level Co-occurence Matrix of Facial Data for Ethnicity Detection", 2018 5th International Multi-Topic ICT Conference (IMTIC), 2018. https://doi.org/10.1109/IMTIC.2018.8467231.
[100] E. Loo, T. Lim, L. Ong and C. Lim, "The influence of ethnicity in facial gender estimation", 2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA), 2018. https://doi.org/10.1109/CSPA.2018.8368710.
[101] Young, D., Sanchez, D. and Wilton, L. (2019). Biracial perception in black and white: How Black and White perceivers respond to phenotype and racial identity cues. https://doi.org/10.1037/cdp0000103.
[102] D. Zhang, Y. Wang and B. Bhanu, "Ethnicity classification based on gait using multi-view fusion", 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 2010. https://doi.org/10.1109/CVPRW.2010.5544614.
[103] L. Krivo, M. Vélez, C. Lyons, J. Phillips and E. Sabbath, "Race, Crime, And The Changing Fortunes of Urban Neighborhoods, 1999–2013", Du Bois Review: Social Science Research on Race, vol. 15, no. 1, pp. 47-68, 2018. https://doi.org/10.1017/S1742058X18000103.
[104] W. Billy Huang, "Immigrant crime in Taiwan: perspectives from Eastern Asia", Foresic Research & Criminology International Journal, vol. 6, no. 3, 2018. https://doi.org/10.15406/frcij.2018.06.00210.
[105] M. Hollis and R. Martínez, "Theoretical Approaches to the Study of Race, Ethnicity, Crime, and Criminal Justice", The Handbook of Race, Ethnicity, Crime, and Justice, pp. 203-207, 2018. https://doi.org/10.1002/9781119113799.part2.
[106] S. Gabbidon and H. Greene, "Race and Crime," SAGE Publication, 2019.
[107] S. Awaworyi Churchill and E. Laryea, "Crime and Ethnic Diversity: Cross-Country Evidence", Crime & Delinquency, vol. 65, no. 2, pp. 239-269, 2017. https://doi.org/10.1177/0011128717732036.
[108] R. Martínez, M. Hollis and J. Stowell, "The Handbook of Race, Ethnicity, Crime, and Justice", 2018. https://doi.org/10.1002/9781119113799.part1.
[109] A. Fernandes and R. Crutchfield, "Race, Crime, and Criminal Justice", Criminology & Public Policy, vol. 17, no. 2, pp. 397-417, 2018. https://doi.org/10.1111/1745-9133.12361.
[110] R. Martínez and M. Hollis, "An Overview of Race, Ethnicity, Crime, and Justice", The Handbook of Race, Ethnicity, Crime, and Justice, pp. 11-16, 2018. https://doi.org/10.1002/9781119113799.part1.
[111] J. Unnever, "Ethnicity and Crime in the Netherlands", International Criminal Justice Review, vol. 29, no. 2, pp. 187-204, 2018. https://doi.org/10.1177/1057567717752218.
[112] T. A. LaVeist, "Beyond dummy variables and sample selection: what health services researchers ought to know about race as a variable," Health Serv Res, pp. 1-16, 1994.
[113] T. Stewart, "Medico-legal aspects of the skeleton. I. Age, sex, race and stature," American Journal of Physical Anthropology, vol. 6, no. 3, pp. 315-322, 1948. https://doi.org/10.1002/ajpa.1330060306.
[114] A. Dubey, S. Roy and S. Verma, "Significance of Sacral Index in Sex Determination and Its Comparative Study in Different Races", 2019. https://doi.org/10.16965/ijar.2016.153.
[115] R. Kumar, "Forensic odontology: A medico legal guide for police personnel", International Journal of Forensic Odontology, vol. 2, no. 2, p. 80, 2017. https://doi.org/10.4103/ijfo.ijfo_13_17.
[116] J. Betancourt, "Defining Cultural Competence: A Practical Framework for Addressing Racial/Ethnic Disparities in Health and Health Care", Public Health Reports, vol. 118, no. 4, pp. 293-302, 2003. https://doi.org/10.1093/phr/118.4.293.
[117] J. Williams, R. Walker and L. Egede, "Achieving Equity in an Evolving Healthcare System: Opportunities and Challenges", The American Journal of the Medical Sciences, vol. 351, no. 1, pp. 33-43, 2016. https://doi.org/10.1016/j.amjms.2015.10.012.
[118] C. Jones, B. Truman, L. Elam-Evans, "Using Socially Assigned Race to Probe White Advantages in Health Status," Race, Ethnicity, and Health: A Public Health Reader, 2013.
[119] R. Harris, D. Cormack and J. Stanley, "The relationship between socially-assigned ethnicity, health and experience of racial discrimination for Māori: analysis of the 2006/07 New Zealand Health Survey", BMC Public Health, vol. 13, no. 1, 2013. https://doi.org/10.1186/1471-2458-13-844.
[120] D. Cormack, R. Harris and J. Stanley, "Investigating the Relationship between Socially-Assigned Ethnicity, Racial Discrimination and Health Advantage in New Zealand", PLoS ONE, vol. 8, no. 12, p. e84039, 2013. https://doi.org/10.1371/journal.pone.0084039.
[121] R. Harris, D. Cormack, J. Stanley and R. Rameka, "Investigating the Relationship between Ethnic Consciousness, Racial Discrimination and Self-Rated Health in New Zealand", PLOS ONE, vol. 10, no. 2, p. e0117343, 2015. https://doi.org/10.1371/journal.pone.0117343.
[122] W. Lee, S. Veeranki, H. Serag, K. Eschbach and K. Smith, "Improving the Collection of Race, Ethnicity, and Language Data to Reduce Healthcare Disparities: A Case Study from an Academic Medical Center," Perspect Health Inf Manag, vol. 13, 2016.
[123] R. Holliday, "Aesthetic surgery as false beauty", Feminist Theory, vol. 7, no. 2, pp. 179-195, 2006. https://doi.org/10.1177/1464700106064418.
[124] T. A. Elsamny, A. N. Rabie, A. N. Abdelhamidand E. A. Sobhi, "Anthropometric Analysis of the External Nose of the Egyptian Males," Aesthetic Plastic Surgery, vol. 42, no. 5, pp. 1343–1356, 2018. https://doi.org/10.1007/s00266-018-1197-8.
[125] I. Mohammed, T. Mokhtari, S. Ijaz, A. Omotosho , A. Ngaski, M. Milanifard and G. Hassanzadeh, "Anthropometric study of nasal index in Hausa ethnic population of northwestern Nigeria," Journal of contemporary medical science, vol. 4, 2018.
[126] A. Sa, O. B, A. Gt, F. Op And I. Ag, "Anthropometric Characterization Of Nasal Parameters In Adults Oyemekun Ethnic Group In Akure Southwest Nigeria", International Journal of Anatomy and Research, vol. 6, no. 22, pp. 5272-5279, 2018. https://dx.doi.org/10.16965/ijar.2018.178.
[127] S. Ozdemir, D. Sigirli, I. Ercan and N. Cankur, "Photographic Facial Soft Tissue Analysis of Healthy Turkish Young Adults: Anthropometric Measurements," Aesthetic Plastic Surgery, vol. 33, no. 2, pp. 175-184, 2008. https://doi.org/10.1007/s00266-008-9274-z.
[128] E. Plemons, "Gender, Ethnicity, and Transgender Embodiment: Interrogating Classification in Facial Feminization Surgery", Body & Society, vol. 25, no. 1, pp. 3-28, 2018. https://doi.org/10.1177/1357034X18812942.
[129] T. Elsamny, A. Rabie, A. Abdelhamid and E. Sobhi, "Anthropometric Analysis of the External Nose of the Egyptian Males", Aesthetic Plastic Surgery, vol. 42, no. 5, pp. 1343-1356, 2018. https://doi.org/10.1007/s00266-018-1197-8.
[130] B. Marsh, K. Pezdek and D. Ozery, "The cross-race effect in face recognition memory by bicultural individuals," Acta Psychologica, vol. 169, pp. 38-44, 2016. https://doi.org/10.1016/j.actpsy.2016.05.003.
[131] T. Tong, T. Key, J. Sobiecki and K. Bradbury, "Anthropometric and physiologic characteristics in white and British Indian vegetarians and nonvegetarians in the UK Biobank", The American Journal of Clinical Nutrition, vol. 107, no. 6, pp. 909-920, 2018. https://doi.org/10.1093/ajcn/nqy042.
[132] N. Mehta and R. Srivastava, "The Indian nose: An anthropometric analysis", Journal of Plastic, Reconstructive & Aesthetic Surgery, vol. 70, no. 10, pp. 1472-1482, 2017. https://doi.org/10.1016/j.bjps.2017.05.042.
[133] E. Ozsahın, E. Kızılkanat, N. Boyan, R. Soames and O. Oguz. "Evaluation of Face Shape in Turkish Individuals," Int. J. Morphol., 2016. https://scielo.conicyt.cl/pdf/ijmorphol/v34n3/art15.pdf.
[134] D. Chettri, D. Sinha, D. Chakraborty and D. Jain, "Naso-facial anthropometric study of Female Sikkimese University Students", IOSR Journal of Dental and Medical Sciences, vol. 16, no. 03, pp. 49-54, 2017. https://doi.org/10.9790/0853-1603074954.
[135] M. Ho and K. Pezdek, "Postencoding cognitive processes in the cross-race effect: Categorization and individuation during face recognition", Psychonomic Bulletin & Review, vol. 23, no. 3, pp. 771-780, 2015. https://doi.org/10.3758/s13423-015-0945-x.
[136] J. Cavazos, E. Noyes and A. O'Toole, "Learning context and the other-race effect: Strategies for improving face recognition", Vision Research, 2018. https://doi.org/10.1016/j.visres.2018.03.003.
[137] J. Anacleto and A. Carvalho, "Improving Human-Computer Interaction by Developing Culture-sensitive Applications based on Common Sense Knowledge," Human-Computer Interaction.
[138] A. Cowell and K. Stanney, "Manipulation of non-verbal interaction style and demographic embodiment to increase anthropomorphic computer character credibility", International Journal of Human-Computer Studies, vol. 62, no. 2, pp. 281-306, 2005. https://doi.org/10.1016/j.ijhcs.2004.11.008.

[139] A. Schlesinger, W. Edwards and R. Grinter, "Intersectional HCI", Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17, 2017. https://doi.org/10.1145/3025453.3025766.

[140] C. Munteanu, H. Molyneaux, W. Moncur, M. Romero, S. O'Donnell and J. Vines, "Situational Ethics", Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15, 2015. https://doi.org/10.1145/2702123.2702481.

[141] A. Schlesinger, W. Edwards and R. Grinter, "Intersectional HCI," Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17, 2017. https://doi.org/10.1145/3025453.3025766.

[142] R. Bellini et al., "Feminist HCI," Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18, 2018. https://doi.org/10.1145/3170427.3185370.

[143] S. Breslin and B. Wadhwa, "Towards a Gender HCI Curriculum," Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15, 2015. https://doi.org/10.1145/2702613.2732923.

[144] S. Yardi and A. Bruckman, "Income, raceand class," Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12, 2012. https://doi.org/10.1145/2212776.

[145] M. Allison and L. Kendrick, “Towards an Expressive Embodied Conversational Agent Utilizing Multi-Ethnicity to Augment Solution Focused Therapy,” International Florida Artificial Intelligence Research Society Conference, pp. 332-337, 2013.

[146] J. Waycott et al., "Ethical Encounters in Human-Computer Interaction", Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16, 2016. https://doi.org/10.1145/2851581.2856498.

[147] P. Kumar, M. Bashour and V. Packiriswamy, "Anthropometric and Anthroposcopic Analysis of Periorbital Features in Malaysian Population: An Inter-racial Study", Facial Plastic Surgery, vol. 34, no. 04, pp. 400-406, 2018. https://doi.org/10.1055/s-0038-1648224.

[148] S. Jilani, H. Ugail and A. Logan, “Inter-Ethnic and Demic-Group Variations in Craniofacial Anthropometry: A Review,” PSM Biological Research, vol. 1, 2019.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 2, 393-398 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Hybrid Autogyro: Model and Longitudinal Control for Wind Gust Energy Conversion Using Autorotation

Jonathan Flores Santiago*[1], Sergio Salazar Cruz[1], Rogelio Lozano Leal[1,2]

[1] *UMI-LAFMIA 3175 CNRS, CINVESTAV, 07360, Mexico.*

[2] *Sorbonne Universités, UTC-CNRS, UMR 7253 Heudiasyc, Compiègne, France.*

ARTICLE INFO

Article history:
Received: 14 January, 2020
Accepted: 11 March, 2020
Online: 07 April, 2020

Keywords:
Autorotation
Longitudinal Control
Hybrid autogyro

ABSTRACT

In the new technologies for producing electricity from wind gusts, a new class of wind power converter has been proposed. The hybrid autogyro uses the principle of autorotation to generate lift. This paper proposes to take advantage of this rotation to also generate energy. There are systems that use flying wings or airplanes to reach winds that blow in layers of the atmosphere that are inaccessible to traditional wind turbines. Wind gusts are disturbances that significantly affect the behavior of this kind of aircraft. However wind gust is the source of energy and this wind speed increases with the aircraft's altitude. This paper reviews the mathematical model of the hybrid autogyro aircraft and proposes a control algorithm to ensure that the vehicle remains horizontal in forward flight, in this way is possible to maximize the generated energy, this concept is tested in numerical simulations.

Symbol	**Description**
a_0, a_1, b_1	Coning, forward, and side flapping angles.
F	Force vector $[X\ Y\ Z]^T$.
g	acceleration due to gravity.
H	Angular momentum vector.
I_{xx}, I_{yy}, I_{zz}	Inertia matrix components.
K_p, K_i, K_d	Gain of control terms.
L, M, N	Torques around x, y and z axes.
m	Autogyro mass.
${}_pN, {}_pE, {}_pD$	Inertial frame positions (north, east, and down).
P, Q, R	Body rotation rates about x, y, and z directions.
$\mathbb{R}$	Rotational matrix, body to inertial frame.
TAF	Total aerodynamic force.
U, V, W	Linear velocities in x, y, and z axes.
v	Velocity vector $[U\ V\ W]^T$.
X, Y, Z	Forces in x, y, and z directions.
δ	Deflection angle of elevator.
τ	Torque vector $[L\ M\ N]^T$.
ν	Noise.
ϕ, θ, ψ	Euler angles (roll, pitch, yaw).
ω	Body rotation rate vector $[P\ Q\ R]^T$.
$\Omega_{3,4}$	Matrices of rotation rates.
$(\cdot)_A$	Aerodynamic.
$(\cdot)_a$	Air relative.
$(\cdot)_d$	Desired angle.
$(\cdot)_w$	Air relative velocities.
$(\cdot)^i$	Inertial frame.
$(\cdot)_e$	Error.
$(\cdot)_{fus}$	Fuselage.
$(\cdot)_{ht}$	Horizontal tail.
$(\cdot)_{pid}$	Proportional, integral and derivative terms.
$(\cdot)_r$	Rotor.
$(\cdot)_{vt}$	Vertical tail.

1 Introduction

The arrival of the 21st century brought a great advance in technology. Unmanned Aerial Vehicles (UAV's) are now the object of interest by sectors of society that find more and more applications, for example to explorer areas which are difficult to access or imply high risk for humans [1]. UAV's present great advances in design and control, due to the versatility of the applications [2, 3]. On the other hand, well-known techniques that use natural resources such as solar radiation, wind gusts or the water flow of rivers, could be improved by using UAV's for inspection and maintenance work [4, 5]. Furthermore, wind generators, solar panels, hydroelectric dams among others, have high-impact environmental and economic costs. In addition, there are cases in which geographical and environmental areas have very specific characteristics [6].

In recent years, air vehicles have presented advances in design and control thanks to a large number of applications and scopes.

*Jonathan Flores Santiago, Gustavo A. Madero, Mexico City 07360, Contact +525523711324, Email: jonathan.flores@cinvestav.mx

One of the advances is to generate energy using wind gusts. In [7] the authors present a study of the technologies used to generate energy. The techniques used for wind gusts can be arranged in three main groups of generation [8]:

- On board.
- Ground generation.
- On board and ground generation.

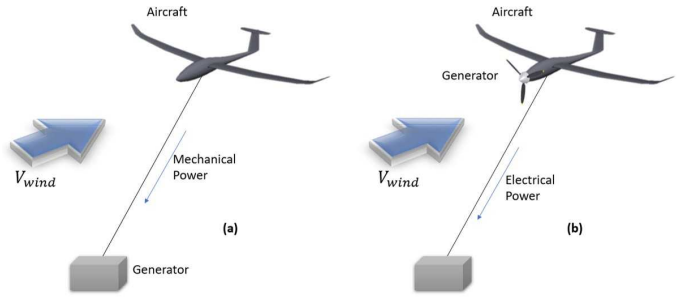


Figure 1: (a) Ground generation. (b) Generation on board.

On board case has the purpose of converting the wind gust to energy using electric generators like the *Makani* Company [9]. Ground generation case, the traction generated by the gust of wind on the aircraft which is attached to some mechanism on the ground. In [10]-[13] are presented different techniques to transform that traction force into energy. This technology is developed to obtain benefits from wind gusts once the aircraft is on the air. However take-off and landing of the aircraft are high risk maneuvers and are performed manually.

The first model of the autogyro was developed by Glauert and Wheatley in the early 30's [14, 15], who introduced the concept of *blade element impulse*, a phenomenon responsible for the state of autorotation. Developed by the Spanish engineer Juan de la Cierva in 1923, the autogyro can be described as the antecedent of the modern helicopter. The principle of self-rotation consists of a free rotation wing for the elevation generated by the induced wind. This aircraft is powered by engine (equivalent in classic airplane). Thanks to self-rotation, landing and take-off are possible on a much shorter runway, even under specific wind conditions, short take-offs or *jump start* is possible. There is a great variety of aircrafts applied to the generation of energy and there are different techniques of flight for obtaining energy.

This paper proposes an autogiro with fuselage and a control surfaces as an airplane to generate pitching moment. The main difference is that it uses a free rotating wing which changes the angle of the lateral inclination to generate the roll moment of the aircraft, similar to the classic helicopter. The implementation of a fixed wing for the autogyro fuselage is presented in [16]-[19], the improvement in performance is shown by the authors.

The paper is organized as follows: Section 2 presents the autorotation in vertical flight and forward flight. Section 3 describes the mathematical model of the autogyro and its dynamics. Section 4 describes the longitudinal control law. Section 5 is devoted to the energy conversion. Final remarks are given in the conclusions in section 6.

2 Autorotation

2.1 Autorotation in vertical flight

Sometimes an auto start system (engine that drives the rotor) is used for the take-off at a very short distance, producing an almost vertical take-off. Similarly, the landing is done almost vertically and in slow-soft landing. The autogyro as others aircrafts obtains the lift from the wind upwards that crosses the disc formed by the rotation of its main rotor inverse to the helicopters. This wind gust generates aerodynamic forces along each propeller which allow the autorotation [20, 21].

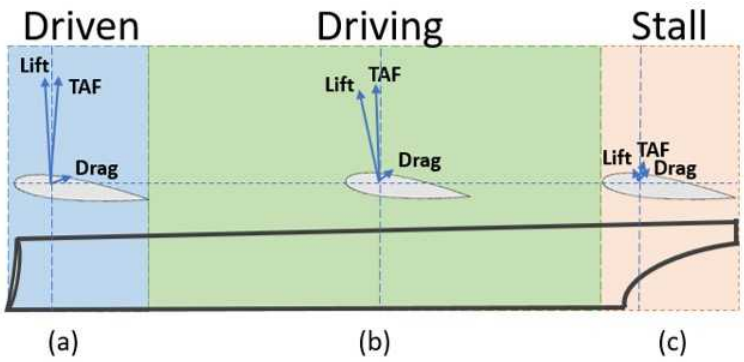


Figure 2: (a) Driven: A bearing force is generated orthogonal to the wind incidence, the total aerodynamic force is soft-opposite to the rotation. (b) Driving: A greater lift force is produced, the region farthest from the center hub is the one with the highest speed, the total aerodynamic force is positive to the rotation. (c) Stall: Mechanically it is where the axis of rotation and hub of the helices is located.

The Lift force, Drag force and the Total Aerodynamic Force or TAF are the forces with which it is possible to generalize its operation and that vary along the regions of each blade [22].

2.2 Autorotation in Forward flight

Landing and take-off are important stages in flight. The forward displacement generates aerodynamic forces that must be considered. The autorotation, takes advantage of the wind incidence. The autogyro must remain in forward flight to generate lift in the same way as an airplane.

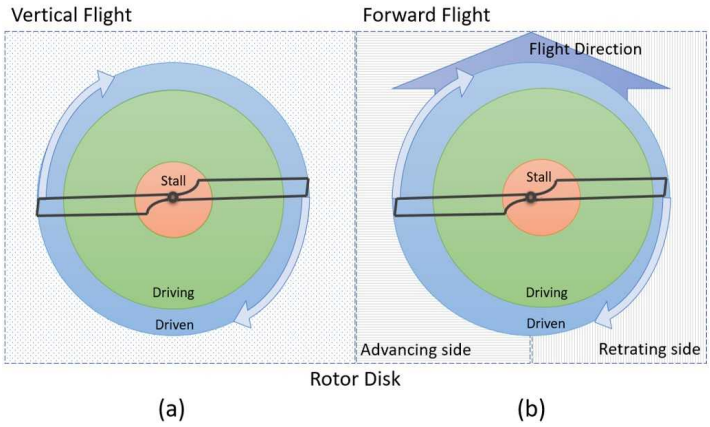


Figure 3: From a top view its possible to separate the rotation disk in left and right hemispheres, therefore the rotation is clockwise.

Figure 3 shows the aerodynamic forces due to the relative wind generated by the forward flight of the autogyro. The forward blade corresponds to the right hemisphere and the retreating blade corresponds left hemisphere. This relative movement of the forward blade presents an increase in its total aerodynamic force which is larger than the one produced by the retracting blade. This difference induces an angular momentum that could cause an overturn [23, 24]. A semi-cyclic plate is able to control the orientation by tilting the rotor disc using the roll moment [25].

3 Mathematical model

3.1 Equations of motion

The six-degree of freedom rigid body dynamics equations will be developed in this section. The fundamental relations between forces and linear momentum rates and torques and angular momentum rates are

$$\begin{aligned} F &= m\dot{v}^i \\ \tau &= \dot{H}^i \end{aligned} \tag{1}$$

where H is the angular momentum vector with the derivatives taken with respect to the inertial frame $*^i$ [26]. The force and torque vectors $F = [X\ Y\ Z]^T$ and $\tau = [L\ M\ N]^T$ are

$$\begin{aligned} F &= m(\dot{v} + \omega \times v) \\ \tau &= \dot{H} + \omega \times H \end{aligned} \tag{2}$$

where the derivatives are taken with respect to the airframe

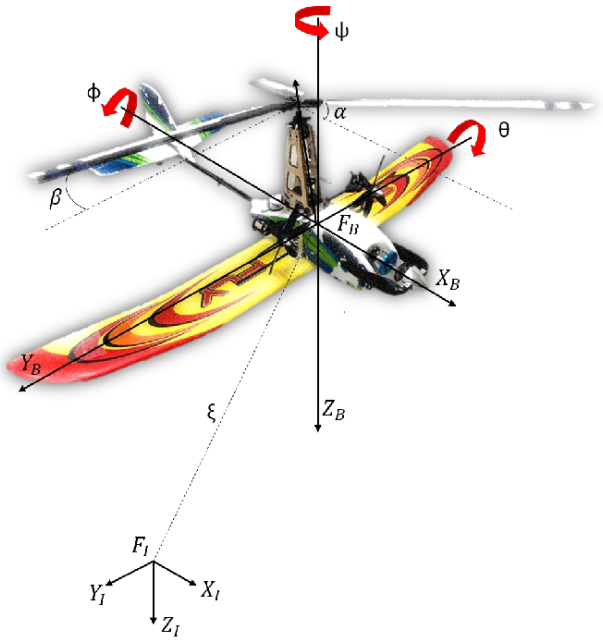


Figure 4: Autogyro Diagram

and $v = [U\ V\ W]^T$ is the relative velocities vector and $\omega = [P\ Q\ R]^T$ is the rotation rates vector. Then, using the matrix cross product, the angular momentum is

$$\omega \times v = \Omega_3 v = \begin{bmatrix} 0 & -R & Q \\ R & 0 & -P \\ -Q & P & 0 \end{bmatrix} \begin{bmatrix} U \\ V \\ W \end{bmatrix} \tag{3}$$

The force and moment equations can be rewritten as

$$\begin{aligned} \tfrac{1}{m} F &= \dot{v} + \Omega_3 v \\ \tau &= I\dot{\omega} + \Omega_3 I\omega \end{aligned}$$

where the individual state derivatives represents the forces and moments vectors, where it is established that in a hybrid autogyro the autorotating propeller is considered as a passive element, therefore the remaining dynamics may well be represented as that of a fixed-wing aircraft according to [27].

$$\begin{bmatrix} \dot{U} \\ \dot{V} \\ \dot{W} \end{bmatrix} = \frac{1}{m} \left(\begin{bmatrix} X_A \\ Y_A \\ Z_A \end{bmatrix} + \begin{bmatrix} X_r \\ Y_r \\ Z_r \end{bmatrix} + C^T \begin{bmatrix} 0 \\ 0 \\ mg \end{bmatrix} \right) + \begin{bmatrix} RV - QW \\ PW - RU \\ QU - PV \end{bmatrix} \tag{4}$$

$$\begin{bmatrix} \dot{P} \\ \dot{Q} \\ \dot{R} \end{bmatrix} = \begin{bmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{bmatrix}^{-1} \left(\begin{bmatrix} L_A \\ M_A \\ N_A \end{bmatrix} + \begin{bmatrix} L_r \\ M_r \\ N_r \end{bmatrix} + \begin{bmatrix} QR(I_{yy} - I_{zz}) \\ RP(I_{zz} - I_{xx}) \\ PQ(I_{xx} - I_{yy}) \end{bmatrix} \right) \tag{5}$$

The weight vector $g = [0\ 0\ mg]^T$ must be included with the direction cosine matrix or rotational matrix $\mathbb{R}$, which rotates vectors from the body to the inertial frame. Euler angles describe the rotation with respect to the inertial frame. $\mathbb{R}$ matrix transform vectors from the body to the inertial frame and is defined as

$$\begin{aligned} \mathbb{R} &= \begin{bmatrix} c_\phi c_\theta & c_\psi s_\theta s_\phi - s_\psi c_\phi & c_\psi s_\theta c_\phi + s_\psi s_\phi \\ s_\phi c_\theta & s_\psi s_\theta s_\phi + c_\psi c_\phi & s_\psi s_\theta c_\phi - c_\psi s_\phi \\ -s_\theta & c_\theta s_\phi & c_\theta c_\phi \end{bmatrix} \\ &= \begin{bmatrix} C_{xx} & C_{xy} & C_{xz} \\ C_{yx} & C_{yy} & C_{yz} \\ C_{zx} & C_{zy} & C_{zz} \end{bmatrix} \end{aligned} \tag{6}$$

where c_ϕ corresponds to $\cos\phi$, s_ϕ corresponds to $\sin\phi$ and the same for $c_{\theta,\psi}$ and $s_{\theta,\psi}$ [28]. The relation between Euler angles and the body angular velocities is the following

$$\begin{bmatrix} P \\ Q \\ R \end{bmatrix} = \begin{bmatrix} \dot{\phi} - \dot{\psi}\sin\theta \\ \dot{\theta}\cos + \dot{\psi}\cos\theta\sin\phi \\ -\dot{\theta}\sin\phi + \dot{\psi}\cos\theta\cos\phi \end{bmatrix} \tag{7}$$

The Euler angles derivatives are

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} P + (Q\sin\phi + R\cos\phi)\tan\theta \\ Q\cos\phi - R\sin\phi \\ (Q\sin\phi + R\cos\phi)\sec\theta \end{bmatrix} \tag{8}$$

Finally, the equations for the position are defined in the local coordinate frame as

$$\begin{bmatrix} {}_p\dot{N} \\ {}_p\dot{E} \\ {}_p\dot{D} \end{bmatrix} = \mathbb{R} \begin{bmatrix} U \\ V \\ W \end{bmatrix} \tag{9}$$

where ${}_pN$, ${}_pE$, ${}_pD$ represent the inertial frame positions (north, east, and down).

3.2 Wind model

The wind is introduced in the local frame, the body velocities $[U\ V\ W]$ are converted to local velocities through the direction cosine matrix $\mathbb{R}$. The wind velocities can be subtracted and the result converted back to the body frame to obtain the air relative velocity

$$\begin{bmatrix} U_a \\ V_a \\ W_a \end{bmatrix} = \mathbb{R}^T \left(\mathbb{R} \begin{bmatrix} U \\ V \\ W \end{bmatrix} - \begin{bmatrix} U_w \\ V_w \\ W_w \end{bmatrix} \right) \tag{10}$$

where $(\cdot)_a$ and $(\cdot)_w$ represent air relative and air relative velocities respectively. For simulation proposes, the wind is generated by the equation

$$v_w = (1 + \nu_{wind}) v_{w_i} G_w \tag{11}$$

where v_w is the constant wind intensity. A zero mean Gaussian noise *turbulence* term ν_{wind} with standard deviation σ_{wind} is added to 1 to perturb v_{w_i} this signal is then shaped by G_w to smooth the turbulence, where G_w is a first order low pass filter.

3.3 Forces and Moments

3.3.1 Rotor Forces and Moments

The forces and moments produced by the rotor considering the distance of the rotor hub from the vehicle center of gravity and if the flapping angles are assumed small, the equations can be simplified with sine and cosine approximations.

$$\begin{bmatrix} X_r \\ Y_r \\ Z_r \end{bmatrix} = \begin{bmatrix} -T_{a_1} \\ T_{b_1} \\ -T \end{bmatrix}$$
$$\begin{bmatrix} L_r \\ M_r \\ N_r \end{bmatrix} = \begin{bmatrix} -T b_1 z_r \\ T_{x_r} - T_{a_1} z_r \\ 0 \end{bmatrix} \tag{12}$$

3.3.2 Aerodynamic Forces and Moments

The total aerodynamic forces and moments in the equations of motion are:

$$\begin{bmatrix} X_A \\ Y_A \\ Z_A \end{bmatrix} = \begin{bmatrix} X_{fus} \\ Y_{fus} + Y_{vt} \\ Z_{fus} + Z_{ht} \end{bmatrix}$$
$$\begin{bmatrix} L_A \\ M_A \\ N_A \end{bmatrix} = \begin{bmatrix} -Y_{fus} z_{fus} - Y_{vt} z_{vt} \\ X_{fus} z_{fus} - Z_{ht} x_{ht} \\ Y_{vt} x_{vt} \end{bmatrix} \tag{13}$$

The approach to estimate these forces and moments is to assume that, similar to a helicopter, the lift force is generated by a rotating wing. However, an important aerodynamic difference is that the wind flow is upward through the rotor. Another difference is the absence of the moment produced by the helicopter motor. It is possible to find a more complete study in [29] where similar details to other aircrafts are considered.

4 Control Structure

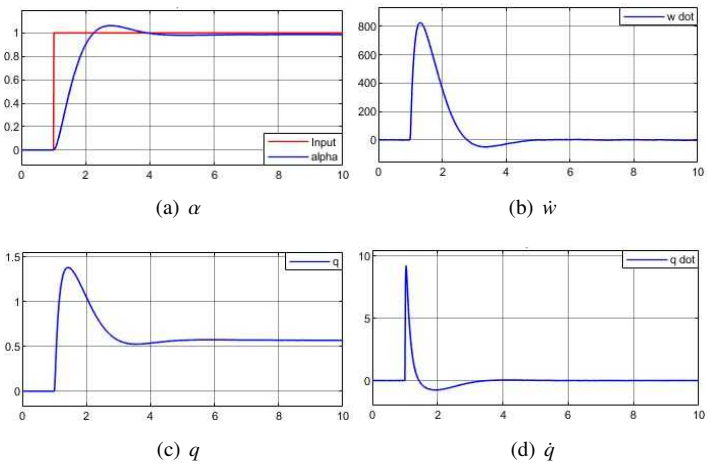
Separate lateral and longitudinal controllers are constructed in decoupled modes and the two can be integrated with minimal risk of interference.

4.1 Longitudinal Control Design

The longitudinal dynamics is governed by a PID controller as an internal loop for attitude control. The controller includes the input pitch angle. The corresponding output is the deflection angles of the elevator δ_e. The longitudinal control law is as follows:

$$\delta_e = \Big[(\theta_d - \theta) K_{p\theta} - p\Big] K_{pq} + K_{iq} \int \Big[(\theta_d - \theta) K_{p\phi} - p\Big] dt \tag{14}$$
$$+ K_{dp} \frac{d\Big[(\theta_d - \theta) K_{p\theta} - p\Big]}{dt}$$

the desired angle is represented by θ_d. The Classical PID control strategy is used to control the attitude of the aircraft [26, 16] due to the fact that the attitude of the hybrid autogyro and the energy generation are independent processes. The PID Tuner app in Matlab was used for automatically tuning the gains of a controller to achieve a balance between performance and robustness.



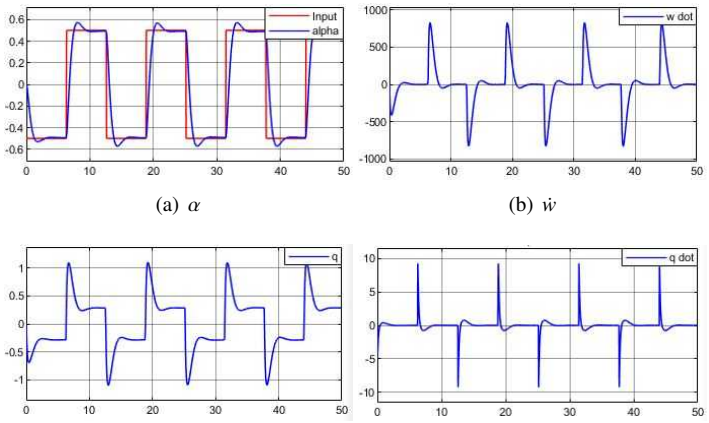
(a) α (b) $\dot{w}$

(c) q (d) $\dot{q}$

Figure 5: Step response. The response of the vehicle to the change of reference.

The inner loop PID controller for the longitudinal control corresponds to the pitch attitude control. The altitude control compares the pitch angle and desired pitch angle and uses it for pitch stabilization to reach the desired altitude (Figure 5). Then it commands the angle of the elevator to control the hybrid autogyro pitch and correct the angle of attack. The longitudinal control of the autogyro commands the flight up and down. In the presence of wind it is necessary to consider the angle of attack of the autogyro which must be regulated to keep the lift at aceptable level. Furthermore, the rotor can be tilted lateraly to make the autogiro turn [30]. To reduce the lifting force loss during the turning and maintain the altitude, this paper adopts the pitching elevator feedforward control method of the propeller (roll moment).

The response of the Hybrid Autogyro to a change of direction presents an effective control before the variation in the reference that may come from the change of direction in the altitude. Figure 6 presents the response of the angle of attack due to a disturbance or a change in altitude, the autogyro reaches its horizontal reference.



(a) α (b) $\dot{w}$

(c) q (d) $\dot{q}$

Figure 6: Square wave signal response.

5 Energy conversion

According to [31] the wind velocity variation can be observed in terms of the altitude. As shown in Figure 7, a terrain with obstacles like buildings or high trees produces a reduction in the wind velocity. In higher levels above ground we can find higher wind velocities but in fact also greater amount of opposite wind gust, which are considered as strong disturbances. The wind velocity variation versus the height above ground exhibits a parabolic shape depending on the ground roughness.

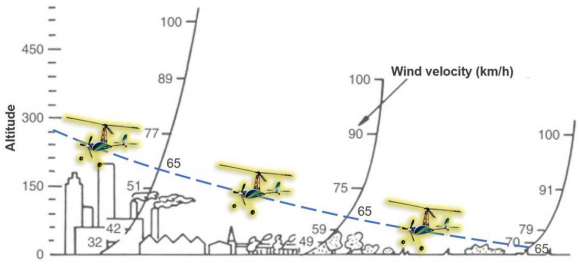


Figure 7: Altitude vs Wind speed for equilibrium flight $65km/h$.(image taken from [31]).

If the speed of the autogyro (induced by the motor) is equal to the minimum wind speed required for the lift, then the autogyro performs a stationary or suspended flight without displacement. Actually, finding this balance in wind speed represents an ideal for energy conversion. The autogyro can remain in flight without using energy in the engine, which represents the highest energy consumption during the flight. Other avionics components (servomotors, sensors, autopilot, etc.), only used for the attitude control, represent a lower energy consumption, which means an increase in flight time [32]. The higher wind speed could represent a negative displacement to the advance of the autogyro, in other words, it would be necessary to anchor the autogyro as a kite.

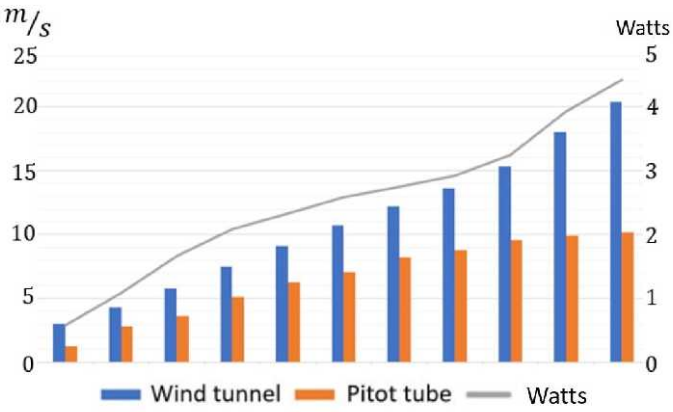


Figure 8: Wind speed vs energy in flight.

When the wind speed is above 65 km/h, the use of the main motor for producing lift may not be necessary. However the longitudinal control is crucial to improve the performance in the generation of energy (Figure 8).

$$V_{wind} \approx 65km/h \rightarrow 5Wh. \quad (15)$$

6 Conclusions

The hybrid autogyro uses the principle of autorotation to generate lift and in this paper has proposed to take advantage of this rotation for energy generation by placing generators on the frame. However, the wind speed should be considerably higher to flight an autogyro than the wind required for fixed-wing vehicles. Therefore, the autorotation serves simultaneously to generate lift and to generate energy. It has been observed that it is important to consider the mathematical model and control algorithm of the autogyro to obtain a good behavior of the aircraft. Consequently and because wind gusts are also non-constant disturbances, the implementation of a control strategy that improves performance in altitude control is important for power generation. Nonetheless the wind is the source for power generation at high altitude.

A wide range of topics for the UAV's that pursue the generation of energy, such as aerodynamics, avionics, control and energy is presented in this paper. Currently these areas are continually being developed promoting clean and renewable energy through specific UAV's projects.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This project was funded by UMI-LAFMIA 3175 CNRS, CINVESTAV and UTC, CNRS, UMR 7253 Heudiasyc. Finally to the National Council of Science and Technology (CONACYT) for the support.

References

[1] Rao, B., Gopi, A. G., & Maione, R. (2016). The societal impact of commercial drones. Technology in Society, 45, 83-90.

[2] Luppicini, R., & So, A. (2016). A technoethical review of commercial drone use in the context of governance, ethics, and privacy. Technology in Society, 46, 109-119.

[3] Finn, R. L., & Wright, D. (2016). Privacy, data protection and ethics for civil drone practice: A survey of industry, regulators and civil society organisations. Computer Law & Security Review, 32(4), 577-586.

[4] Rakha, T., & Gorodetsky, A. (2018). Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones. Automation in Construction, 93, 252-264.

[5] Jo, J., Jadidi, Z., & Stantic, B. (2017, October). A Drone-Based Building Inspection System Using Software-Agents. In International Symposium on Intelligent and Distributed Computing (pp. 115-121). Springer, Cham.

[6] Del Río, P. (2010). Analysing the interactions between renewable energy promotion and energy efficiency support schemes: The impact of different instruments and design elements. Energy Policy, 38(9), 4978-4989.

[7] Cherubini A, Papini A, Vertechy R, Fontana M. (2015). Airborne Wind Energy Systems: A review of the technologies. Renewable and Sustainable Energy Reviews, 51, 1461-1476.

[8] Breukelman, P., Kruijff, M., Fujii, H. A., Maruyama, Y. (2014). A new wind-power generation method employed with high altitude wind. In International Conference and Exhibition, Grand Renewable Energy, Tokyo July.

[9] https://makanipower.com

[10] Ahmed M, Hably A, Bacha S. (2013). Kite generator system modeling and grid integration. IEEE Transactions on Sustainable Energy, 4(4), 968-976.

[11] White FM, Wolf DF (1968). A theory of three-dimensional parachute dynamic stability. Journal of aircraft, 5(1), 86-92.

[12] Argatov I, Rautakorpi P, Silvennoinen R. (2009). Estimation of the mechanical energy output of the kite wind generator. Renewable Energy, 34(6), 1525-1532.

[13] Canale M, Fagiano L, Milanese M. (2009). KiteGen: A revolution in wind energy generation. Energy, 34(3), 355-361.

[14] Wheatley JB. (1935). An Aerodynamic analysis of the autogiro rotor with a comparison between calculated and experimental results.

[15] Glauert H. (1926). A general theory of the autogyro (Vol. 1111). HM Stationery Office.

[16] Song, Z., Cai, Z., Li, K., Zhao, J., Liu, N. (2017, June). Model and Attitude Control of a Miniature Hybrid Autogyro. In Chinese Intelligent Automation Conference (pp. 751-760). Springer, Singapore.

[17] Liu, N., Cai, Z., Ma, Y., Wang, Y. (2017). Control and Flight Testing of a Miniature Compound Autogyro. In AIAA Atmospheric Flight Mechanics Conference (p. 3891).

[18] Hou, J., Cai, Z., Zhao, J., Wang, Y., Liu, N., Song, Z. (2017, December). Fuzzy and PSO-based PID controller for a hybrid autogyro attitude control. In Control Conference (ASCC), 2017 11th Asian (pp. 2160-2165). IEEE.

[19] Mackertich S, Das T. (2016, July). A quantitative energy and systems analysis framework for airborne wind energy conversion using autorotation. In American Control Conference (ACC), 2016 (pp. 4996-5001). IEEE.

[20] Duda, H., Pruter, I. (2012). Flight performance of lightweight gyroplanes. In 28th International Congress of the Aeronautical Sciences, Brisbane, Australia.

[21] Handbook, F. A. A. (2000). 8083-21. Rotorcraft Flying Handbook.

[22] Lin, Q., Cai, Z., Wang, Y. (2014, August). Design, model and attitude control of a model-scaled gyroplane. In Guidance, Navigation and Control Conference (CGNCC), 2014 IEEE Chinese (pp. 1282-1287). IEEE.

[23] Wheatley, J. B. (1934). The aerodynamic analysis of the gyroplane rotating-wing system.

[24] Lopez, C. A., Wells, V. L. (2004). Dynamics and stability of an autorotating rotor/wing unmanned aircraft. Journal of guidance, control and dynamics, 27(2), 258-270.

[25] Houston, S. S. (1998). Identification of autogyro longitudinal stability and control characteristics. Journal of Guidance, Control, and Dynamics, 21(3), 391-399.

[26] Deyst, J. J., & George, S. (2005). Simulation and control design of a gliding autogyro for precision airdrop (Doctoral dissertation, Massachusetts Institute of Technology).

[27] Arnold, U. Investigation of gyroplane flight mechanics (in German). Master Theses, Institute of Flight Systems, Braunschweig, 1988.

[28] Lozano, R. (Ed.). (2013). Unmanned aerial vehicles: Embedded control. John Wiley and Sons.

[29] D. H. Lorenz Meier and M. Pollefeys, " px4: A node-based multithreaded open source robotics framework for deeply embedded platforms," International Conference on Robotics and Automation (ICRA), 2015.

[30] Breukelman, P., Kruijff, M., Fujii, H. A., Maruyama, Y. (2014). A new wind-power generation method employed with high altitude wind. In International Conference and Exhibition, Grand Renewable Energy, Tokyo July.

[31] Hartmann, D. L. (2015). Global physical climatology (Vol. 103). Newnes.

[32] Flores, J., Salazar, S., & Lozano, R. (2019, June). Hybrid autogyro: Airborne wind gust energy conversion using autorrotation. In 2019 International Conference on Unmanned Aircraft Systems (ICUAS) (pp. 1255-1260). IEEE.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 399-410 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Using Metaheuristics-Based Methods to Provide Sustainable Market Solutions, Suitable to Consumer Needs

Ricardo Simões Santos[*,1], António João Pina da Costa Feliciano Abreu[2], Joaquim José Rodrigues Monteiro[3]

[1]*GOVCOPP Universidade de Aveiro, ISEL, Polytechnic Institute of Lisbon, 1959-007, Portugal*

[2]*CTS-UNINOVA Universidade Nova de Lisboa, ISEL, Polytechnic Institute of Lisbon, 1959-007, Portugal*

[3]*INESC-ID Lisboa, ISEL, Polytechnic Institute of Lisbon, Portugal, 1959-007, Portugal*

ARTICLE INFO

Article history:
Received: 01 January, 2020
Accepted: 19 March, 2020
Online: 04 April, 2020

Keywords:
Sustainability
Energy efficiency
Household Appliances
Multi-criteria
Multi-objective Optimization
Evolutionary Algorithms
Simplex

ABSTRACT

Nowadays the buildings sector is one of the key sectors to achieve sustainability, with the correspondent consumers, in particular, the household consumers, having the need to perform sustainable choices every day, regarding the appliances to be acquired from the market. This is not only due to government's growing concerns about sustainability but also with the consumers on having sustainable solutions, given the different economic, social (including their comfort) and environmental needs. However, the existence of several electrical appliances on market, with all their different issues, brands and models, together with the several tradeoffs referred before, difficult the consumer's choices, on having sustainable solutions in the market. Therefore, this work, presents an approach, by using Multi-Attribute Value Theory (MAVT), integrated with metaheuristics, which uses Evolutionary Algorithms (EA) to provide suitable and sustainable market solutions to a consumer, according to its own needs. Based on the achieved solutions and considering the relative importance, given to each consumer, and regarding each dimension of sustainability, it's possible to achieve several savings, namely electrical and water consumption, CO_2 emissions, among others. A case study shall be shown, to demonstrate the applicability of the proposed approach.

1. Introduction

This paper is an extension of work [1] originally presented in 2019 International Young Engineers Forum (YEF-ECE), held in Nova University, Costa da Caparica, Lisbon, Portugal.

Sustainable measures are relevant to reach environmental and economic sustainability, particularly regarding the reduction of energy production dependency from fossil fuels, where buildings account for about 30–43 % of the final energy consumed [2-3].

Based on [4], and from the percentage referred above, the household's sector represents approximately 18 % of the final electric energy consumed in the world, which represents an important sector to be improved in terms of energy efficiency, and through the adoption of sustainable appliances.

According to [5-7], there were some improvements regarding energy efficiency, and related to electrical appliances, by establishing reglementary labeling measures, to provide relevant information to the consumer.

Some of these issues include energy consumption, machine's noise, refrigerator capacity, water consumption, among others [7].

According to [7], such measures, have been arisen, not only in the European countries but as well as in other world regions, such as Africa, America and Asia.

Although the importance of such measures to reach sustainability, there are too much options available in the market, which makes it difficult to know what's the best solution to choose, in order to attend the consumer's needs [6-7].

The difficulty raises, when each appliance has its own set of features, which can be different according to the appliance's brand and model to acquire [4-5].

Considering the diversity of options from the market, and related to each energy service (e.g. lighting, air conditioners, refrigerators, etc.), the number of possible combinations could exponentially rise.

* Ricardo Santos, Email: ricardosimoessantos84@gmail.com

The use of optimization algorithms, to assist the consumer on its decisions, allows the achievement of efficient and even sustainable solutions, regarding the household appliances to acquire.

However, and due the high combinatorial nature of the problem referred before, the use of traditional optimization techniques, could present some disadvantages, regarding the efficiency on reaching a feasible solution, as well as with the number of potential and feasible solutions available to the consumer [8]. This could happen, since that some optimization techniques, based on gradient methods, could be stuck into to a local maxima or minima, limiting therefore, the exploitation of the entire feasible region [8-9].

An alternative way to surpass such disadvantages, could pass by using Evolutionary Algorithms (EA), given the widely use, reflected on the number of works from the literature, on solving optimization problems in less time than other traditional optimization methods (e.g. [8-12]).

Other approaches have been used by the authors, in order to develop an integrated approach that allows to achieve sustainable solutions.

In [1][13-14] the authors have started to use a single objective approach to achieve efficient solutions from the market of household appliances.

The successful contribution to the main goal of their research, have led them to adapt the first approach into a Multiobjective one, given the several dimensions of sustainability (referred before) to attend, and considering each household appliance, related to each energy service to be acquired, namely; Economic, Social and Environment.

In order to consider other kind of consumer's issues, such as its social preferences (e.g. reliability, design, performance, among others) we've applied multiple-attribute value theory (MAVT). This theory allows to model the consumer's choices, as well as other preferences regarding the environment (e.g. CO2 savings, water savings, noise, among others) and the economic dimension (e.g. initial investment, noise, water consumption, among others) [15].

All the dimensions referred above have some constraints, suitable to each case (e.g. number of building occupants, area of the room to be climatized, the type of division, among others), such as the ones related with human comfort (e.g. minimum illuminance to achieve visual comfort, noise, minimum air conditioner to achieve heating comfort), with economy (e.g. budget, water consumption), and environmental (e.g. water savings, CO_2 savings, among others). To maximize each dimension, modelled according to MAVT, it was used optimization techniques to get sustainable solutions from the market to the consumer, by acting on three dimensions (or objective functions) referred before.

According to [16], Evolutionary Algorithms (EA), have been applied with success, to solve many optimization problems within less time, when compared with other methods. However, such methods, uses too many control parameters, which makes them, quite sensitive to the input values, therefore, the need to adjust them [16-17], by preforming some robustness and sensitive tests.

The aim of this research, is to propose an approach, where MAVT is integrated with Non-Dominated Sorting Genetic Algorithm II (NSGAII) [18], based on genetic algorithms (GAs), in order to provide the consumer with sustainable solutions from the market, that not only attends it needs, but as well the tree dimensions referred before; Economic, Social and Environmental. To pursue this goal, we've tested, validated and presented in [1], a first approach, which has consisted in one objective function regarding the economic dimension, where it was tried different problem formulations, to study their influence in terms of results.

Based on the works presented on [1] we've selected the objective function with the best results and tested the entire model's robustness, regarding the influence of GAs parameters. Some results are presented with this preliminary approach.

The 2nd approach presented here, will integrate the developments achieved with the 1st one, into a Multiobjective model by using NSGAII optimization method, where not only economic dimension will be included, but also the environmental one as well. For this purpose, new attributes will be added to the model, regarding the environmental dimension, whose diversity and the correspondent units involved, will be integrated into a unique model by recurring to MAVT.

Finally, it will be obtained a combined solution that attends the economic and environmental dimensions, in order to pursue the main goal referred before. A case study will be presented here, to illustrate the applicability of the method, by presenting an example of a set of a feasible and sustainable solutions, suitable to the consumer needs.

2. Literature review

Several methods such as scenario's analysis (e.g. [19]), are frequently used to simulate a constrained set of solutions.

However, some methods are essentially economical, by allowing the consumers to obtain highest values of energy savings, given the same value of investment (e.g. [17]). Other methods, existed on literature, allows to deal with different issues (e.g. savings with Greenhouse Gas (GHG) emissions, benefit-cost analysis, initial investment costs, among others). Such methods are mainly related to retrofitting measures (e.g. [20]), where some of them are even combined with technologies too (e.g. [21]).

However, there is some limits with these approaches, since they don't consider other relevant factors, such as the environment, the labelling system and legal and social issues as well, to find suitable solutions, to fulfill the requirements of the occupants of the building. These methods also don't consider the attributes, regarding each electrical appliance, which varies according to the number of building's occupants.

Recently, some works are based on multi-criteria decision-making (MCDM) methods, in order to help a consumer to solve his problems regarding the measures of retrofitting with their own building's, by accounting not only energy efficiency factors, but also assuring the comfort of the building's occupants (e.g. [15]).

Regarding other works from the literature, there is other MCDM models, based on MAVT approaches, allowing therefore, the integration of optimization with multicriteria methods, which allows the achievement of feasible solutions, selected based on a set of attributes, organized according to a set of criteria (e.g. [15]).

However, these methods don't account the different attributes, related to each electrical appliance existed on market and adjusted to the individual needs of the occupants.

The use of Metaheuristics to achieve from the energy problems' solving, a set of feasible solutions (e.g.[8], by using particle swarm optimization, e.g. [11], by using genetic algorithms, among others existed on literature), have been increased in the last years. However, none of these approaches, have been considered

into a combined method, that allows the consumer, to pick from the market, a set of sustainable appliances, based on its preferences.

3. Objectives

In order to fulfill the gap referred on literature review, the main objective of this research work is to present an approach to assist an household consumer, who wants to buy sustainable appliances, existed on market, for its home, that not only allows to fulfill its needs, but also allows to accomplish a set of requirements regarding sustainability, based on the three dimensions of sustainability referred before, namely, the Consumer's economic well-being, social well-being and Consumer's environment well-being.

Additionally, it will be presented two approach, which allows to obtain several and alternative solutions from the market, facing therefore, some contingencies that eventually may occur, namely the situations of "out of stock" for instance, given an electrical appliance initially recommended by this method.

The approach developed here, also considers economic, social and environment constraints, regarding each energy service considered in this work.

4. Research Method

4.1. Problem description and case study

The problem presented in this work has considered a consumer, who is intended to acquire a set of household appliances, available on market.

The pre-criteria, considered by the algorithm to pre-select a set of appliances existed on market (Table 1), were instantiated according to the number of occupants, existed in the building. In this case study, it was considered a family (which includes the decision-agent) of 4 occupants.

Table 1: Criteria used

Criteria used	Appliance	Quantity
• Air conditioners (types) considered: • Zone to be heated/ cooled by the air conditioner • Capacity (minimum)	**Air Conditioner**	• Mono split Multisplit, Portable • Living room • 9923,7 BTU
• Capacity, based on the number of household's occupants	**Washing Machine**	• 5 kg
• Load capacity.	**Dishwasher**	• 10 Cutleries
• Useful volume, available for cooking based on the number of occupants	**Oven**	• 44,2 cm x 61,8 cm
• Type of dryer machines • Load capacity	**Dryer machine**	• By condensation ; • 6 kg
• Technology	**Lighting**	• Halogen • CFL • Fluorescent
• Capacity of the fridge • Type of fridge, based on the number of occupants	**Refrigerator**	• 120 liters • Refrigerator Combined type.

However, the values regarding each attribute, can be modified, based on the number of occupants, existed in the building.

Thus, the appliances, was pre-selected from the market, by using the criteria presented on Table 1, in order to reduce the decision space, considering therefore, only the suitable solutions to the consumer needs. This was done, in order to increase EA's efficiency, regarding both approaches, by achieving optimal solutions with less time.

4.2. 1st Approach – Single objective model

The 1st approach, proposed in this work, can be seen on Figure1, where each consumer's option (x_{ij}), regards to an existent solution from the market, which can be defined as a choice *i*, belonged to a certain type of appliance (energy service) *j*, to be acquired by the consumer. Considering the trade-offs referred before, together with the diversity of features, related to each solution, the consumer will deal with a problem of combinatorial nature, whose number of combinations are dependent on the number of options to be considered, regarding each dimension. The 20 million combinations (approximately) considered in this work, can be reduced, by considering that the consumer cannot perform any choices (x_{ij}), given its limited budget.

Furthermore, and based on Table 1, all the equipment's, were pre-selected according to the number of occupants, in order to meet the consumer requirements (e.g. washing machine capacity, capacity of the fridge, etc.). Issues, such as air conditioner minimum capacity, minimum illuminance among others, was also considered here and compiled in a set of criteria (Figure1).

Additionally, it was also considered the influence of the consumer usage profile, regarding each potential solution/option x_{ij}, based on the assumptions presented on Table 2, for the case study considered.

Such assumptions have included the number of hours that each household appliance will perform for each day, which was then extrapolated for a monthly and yearly basis.

Table 2: Assumptions, according to consumer usage profile

Emission factor [gCO_2/kWh]	674,6	**Discount Factor [%]**		7,0
Life cycle (usage phase) considered [years]:	11	**Annual Factor**		7,1
Electrical Energy tariff [€/kWh]	0,171			
Electrical appliance	**Profile of Usage (h)**			
	Day	**Week**	**Month**	**Year**
Air Conditioning	3	14	60	719
Dryer Machine	1,2	8,4	32	361
(Electric) Oven	7	7	31	360
Refrigerator	11	76	333	3960
Lighting	3,1	21	94	1080
Dish Washer Machine	1,1	7,7	32	360
Washer Machine	1,5	10,5	31	358

These parameters have some influence on consumption, regarding each household appliance. Therefore, their savings, was achieved through the comparison of each efficient solution with the correspondent standard one. The consumer can also define its

own profile of usage, according to its needs, or by using the profile, considered in this work, by default.

Regarding each individual efficient solution (x_{ij}), it was determined the correspondent savings, $S_{cons_{i,j}}(x_{ij})$ and $S_{inv_{i,j}}(x_{ij})$, respectively for the consumption and investment and regarding each option *i*, belonging to an appliance type *j*. Both savings, were obtained, based on the comparison between the efficient values of consumption and investment, and the correspondent standard solution (i.e. less efficient), regarding to an appliance/energy service type *j*. The value, resulted from this difference, was then discounted, according to the life cycle period (usage phase) considered, by applying a discount factor (Table 2).

Based on the approach, shown on Figure1, the decision variables are:

$$x_{ij} : j(appliance\,type) = \{1..7\} \wedge i(options) = \{1..n_i\} \quad (1)$$

With the objective to maximize $V_R(x_{ij})$, i.e.:

$$\max\ V_R(x_{ij}) : V_R \in \mathbb{R} \wedge x_{ij} \in \{0,1\} \quad (2)$$

Based on work of [22], the value of V_R, can be achieved by:

$$V_R^{(n)} = \sum_{i=1}^{20}\left(\frac{P_{cons_{i,1}}(x_{ij})}{P_{inv_{ij}}(x_{ij})}\right).\frac{I_{inv.ef_{ij}}(x_{ij})}{I_{stdj}}.x_{i1} + \sum_{i=1}^{11}\left(\frac{P_{cons_{i,2}}(x_{i2})}{P_{inv_{i,2}}(x_{i2})}\right).\frac{I_{inv.ef_{ij}}(x_{i2})}{I_{total}}.x_{i2} + \sum_{j=3}^{7}\sum_{i=1}^{10}\left(\frac{P_{cons_{i,j}}(x_{ij})}{P_{inv_{i,j}}(x_{ij})}\right).\frac{I_{inv.ef_{ij}}(x_{ij})}{I_{stdj}}.x_{ij} \quad (3)$$

Given that only one option *i,* can be chosen, regarding each appliance, the correspondent constraints are:

$$R_1(x_{i1}) : \sum_{i=1}^{20} x_{i1} = 1 \wedge x_{i1} \in \mathbb{R}, \forall i = \{1..20\} \wedge x_{i1} = \{0,1\} \quad (4)$$

$$R_2(x_{i2}) : \sum_{i=1}^{11} x_{i2} = 1 \wedge x_{i2} \in \mathbb{R}, \forall i = \{1..11\} \wedge x_{i2} = \{0,1\} \quad (5)$$

$$R_k(x_{ij}) : \sum_{i=1}^{10} x_{ij} = 1 \wedge k = \{3..7\} \wedge x_{ij} = \{0,1\}, \forall i = \{1..10\}, j = \{3..7\} \quad (6)$$

The budget constraint, can be modeled by using the total investment (budget) and according to the following expression:

$$R_8(x_{ij}) : I_{total}(x_{ij}) = \sum_{i=1}^{20} x_{i1} * I_{ef_{i1}} + \sum_{i=1}^{11} x_{i2} * I_{ef_{i2}} + \sum_{j=3}^{7}\sum_{i=1}^{10} x_{ij} * I_{ef_{ij}} \quad I_{total} \in \mathbb{R} \quad (7)$$

On Figure2, it's presented the EA's individual framework.

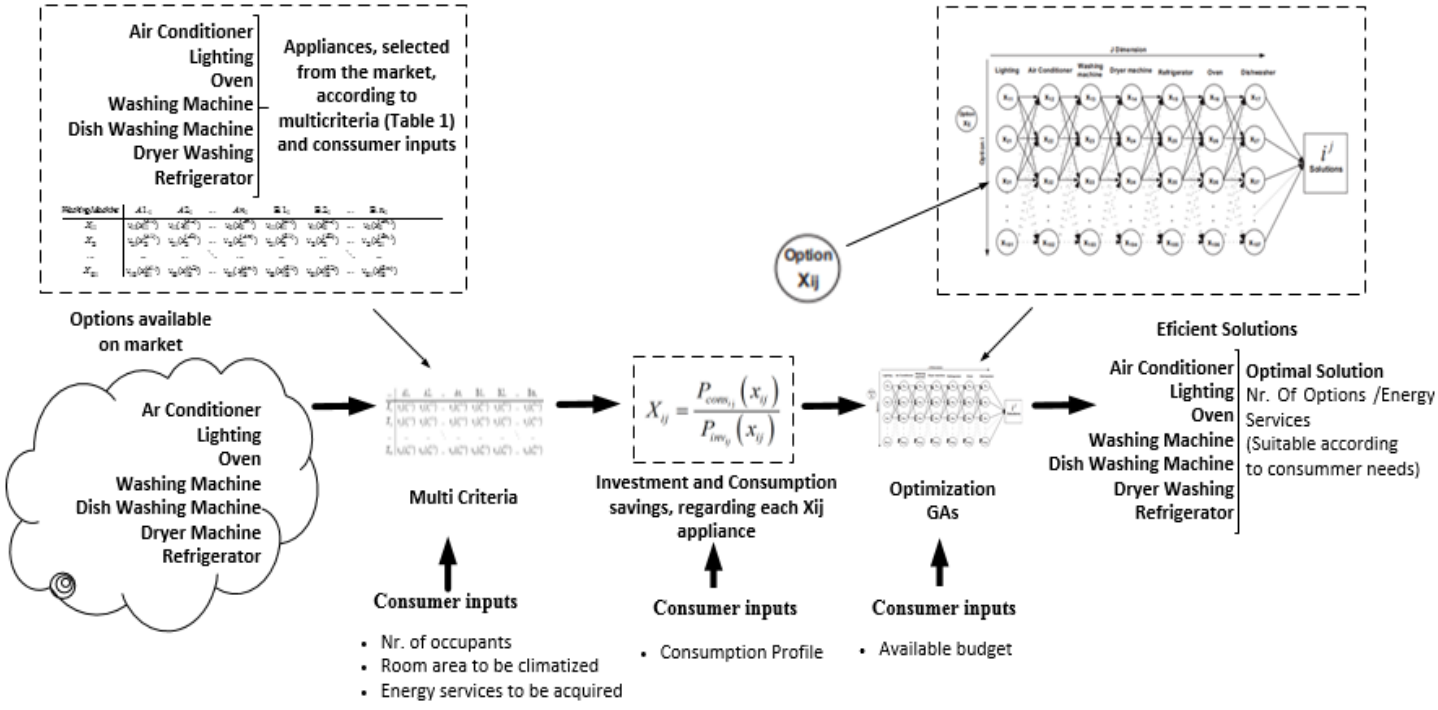


Figure 1(a): 1st Approach – Single objective approach

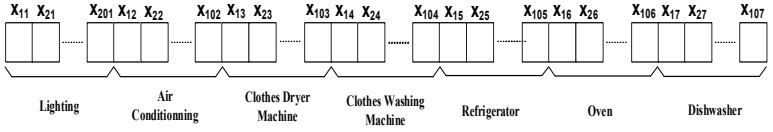


Figure 1(b): EA's individual framework (single objective approach)

Table 3. Definition of the problem dimensions, according to each energy service considered and criteria used

Energy Service	Dimension A-Economics	Ref.	Dimension B-Social	Ref.	Dimension C-Environment	Ref.
Ilu—lighting	Energy Efficiency Classification	Ilu.A1	Durability (h)	Ilu.B1	CO_2e Emissions (Life Cycle—Usage Phase) [kg]	Ilu.C1
	...	...	...	...	...	...
	Energy Cons. Savings (Life Cycle—Usage phase) [€]	Ilu.A5	Color Rendering Index (CRI) [%]	Ilu.B5	CO_2e Emissions (Life Cycle—Production Phase) [kg]	Ilu.C3
AC—Air Conditioner	Energy Efficiency Classification (Heating)	AC.A1	Noise (Indoor) [dB]	AC.B1	CO_2e Emissions (Life Cycle—Production Phase) [kg]	AC.C1
	...	...	...	...	...	...
	Energy Efficiency labelling (Cooling)	AC.A6	Customer Service (warrant)	AC.B9	CO_2e Emissions (Life Cycle—Usage Phase) [kg]	AC.C3
...	...	...	...	...	...	...
FE.—Oven (Electric)	Energy Efficiency Classification	FE.A.1	CO_2e Emissions (Life Cycle—End Use Phase) [kg]	FE.B1	CO_2e Emissions (Life Cycle—Usage Phase) [kg]	FE.C.1
	...	...	...	...	...	...
	Investment cost [€]	FE.A.5	CO_2e Emissions (Life Cycle—End Use Phase) [kg]	FE.B.5	CO_2e Emissions (Life Cycle—End Use Phase) [kg]	FE.C.3
MLL—Dishwasher	Energy Efficiency Classification	MLL.A.1	Water Consumption (Life Cycle—Usage phase)	MLL.B.1	CO_2e Emissions (Life Cycle—Usage phase) [l]	MLL.C.1
	...	...	...	...	...	...
	Water Cons. Savings (Life Cycle—Usage phase) [€]	MLL.A.6	Customer Service (warrant)	MLL.B.6	Water Consumption (Life Cycle—Usage phase) [l]	MLL.C.4

4.3. 2nd Approach – Multiobjective model involving 2 dimensions

The 2nd Approach, was pretended to study the interaction between 2 of the three existed dimensions, regarding sustainability, namely Economics vs Social Wellbeing and Economics vs Environment Wellbeing. To do this, it was included the formulation presented and studied before, regarding economic wellbeing, and the two other dimensions, by recurring to Multiatribute Value Theory (MAVT). This was preformed, in order to convert the objective function (3) (Economic Wellbeing) into a correspondent one, without units involved, to be added further into the other dimensions (Social and/or Environmental). After, the tests performed with multiobjective approach (2 dimensions), we have combined the three objective functions, regarding the three dimensions referred before, into one objective function, to achieve sustainable results, regarding the three dimensions. The assumptions presented before, were the same as well as the case study to apply the approach.

The criteria adopted here, are also the same, although there is the possibility of being changed, based on the number of occupants, existed in the building.

As it referred before, such pre-selection allows to reduction of the decision space, by considering only the available solutions, adjusted to the needs of the consumer. Such pre-selection, contributes also to rise the efficiency of NSGAII, by getting optimal solutions within less time. To pursue this, the proposed approach, has been developed by starting to consider at first 2 and then, the 3 dimensions of the problem. Both cases, are based on the approach described on Figure 2.

Each potential solution (x_{ij}), is regarded from an option *i*, which is correspondent to the energy service *j* to be bought by the consumer from the market. Like the first approach, the consumer's consumption profile (Table 2), was considered, in order to preform Life Cycle Cost Assessment (LCCA) to achieve for each appliance, the corresponding savings, in terms of water consumption ($S_{H2O.Cons_{i,j}}(x_{ij})$), energy consumption ($S_{E.Cons_{i,j}}(x_{ij})$) and initial investment ($S_{inv_{i,j}}(x_{ij})$). All these savings, have been obtained from the comparison between the more efficient solution and the less efficient one ("standard" solution).

By considering several factors, regarding each appliance, as well as the consumer's social, environmental and economic

concerns, a set of criteria was established based on the consumer preferences,

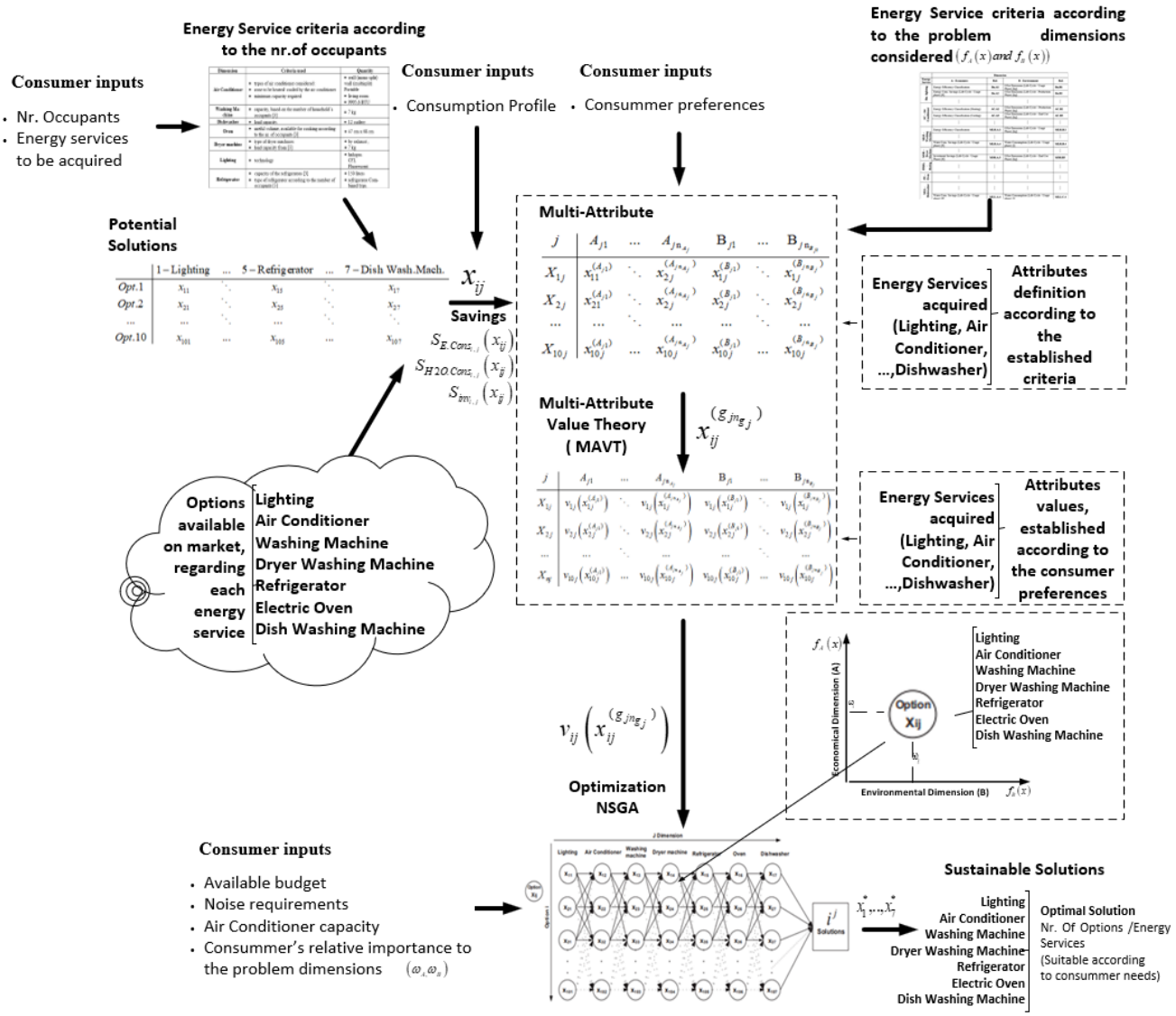


Figure 2: 2nd Proposed approach – Multiobjective approach with MAVT

A.C.	A_{21}	A_{22}	...	$A_{2n_{A_2}}$	B_{21}	B_{22}	...	$B_{2n_{B_2}}$	C_{21}	C_{22}	...	$C_{2n_{C_2}}$
X_{12}	$x_{12}^{(A_{21})}$	$x_{12}^{(A_{22})}$	...	$x_{12}^{(A_{2n_{A_2}})}$	$x_{12}^{(B_{21})}$	$x_{12}^{(B_{22})}$	...	$x_{12}^{(B_{2n_{B_2}})}$	$x_{12}^{(C_{21})}$	$x_{12}^{(C_{22})}$	...	$x_{12}^{(C_{2n_{C_2}})}$
X_{22}	$x_{22}^{(A_{21})}$	$x_{22}^{(A_{22})}$	...	$x_{22}^{(A_{2n_{A_2}})}$	$x_{22}^{(B_{21})}$	$x_{22}^{(B_{22})}$	...	$x_{22}^{(B_{2n_{B_2}})}$	$x_{22}^{(C_{21})}$	$x_{22}^{(C_{22})}$	...	$x_{22}^{(C_{2n_{C_2}})}$
...	...	...	⋱	...	...	...	...	...	...	...	...	...
X_{102}	$x_{102}^{(A_{21})}$	$x_{102}^{(A_{22})}$	...	$x_{102}^{(A_{2n_{A_2}})}$	$x_{102}^{(B_{21})}$	$x_{102}^{(B_{22})}$	...	$x_{102}^{(B_{2n_{B_2}})}$	$x_{102}^{(C_{21})}$	$x_{102}^{(C_{22})}$	...	$x_{102}^{(C_{2n_{C_2}})}$

a)

A.C.	A_{21}	A_{22}	...	$A_{2n_{A_2}}$	B_{21}	B_{22}	...	$B_{2n_{B_2}}$	C_{21}	C_{22}	...	$C_{2n_{C_2}}$
X_{12}	$v_{12}(x_{12}^{(A_{21})})$	$v_{12}(x_{12}^{(A_{22})})$	...	$v_{12}(x_{12}^{(A_{2n_{A_2}})})$	$v_{12}(x_{12}^{(B_{21})})$	$v_{12}(x_{12}^{(B_{22})})$	...	$v_{12}(x_{12}^{(B_{2n_{B_2}})})$	$v_{12}(x_{12}^{(C_{21})})$	$v_{12}(x_{12}^{(C_{22})})$	...	$v_{12}(x_{12}^{(C_{2n_{C_2}})})$
X_{22}	$v_{22}(x_{22}^{(A_{21})})$	$v_{22}(x_{22}^{(A_{22})})$	...	$v_{22}(x_{22}^{(A_{2n_{A_2}})})$	$v_{22}(x_{22}^{(B_{21})})$	$v_{22}(x_{22}^{(B_{22})})$	...	$v_{22}(x_{22}^{(B_{2n_{B_2}})})$	$v_{22}(x_{22}^{(C_{21})})$	$v_{22}(x_{22}^{(C_{22})})$	...	$v_{22}(x_{22}^{(C_{2n_{C_2}})})$
...	...	...	⋱	...	...	...	...	...	...	...	...	...
X_{102}	$v_{102}(x_{102}^{(A_{21})})$	$v_{102}(x_{102}^{(A_{22})})$	...	$v_{102}(x_{102}^{(A_{2n_{A_2}})})$	$v_{102}(x_{102}^{(B_{21})})$	$v_{102}(x_{102}^{(B_{22})})$	...	$v_{102}(x_{102}^{(B_{2n_{B_2}})})$	$v_{102}(x_{102}^{(B_{21})})$	$v_{102}(x_{102}^{(C_{22})})$	...	$v_{102}(x_{102}^{(C_{2n_{C_2}})})$

b)

Figure 3: Example of evaluation table regarding the Air Conditioning appliance's type: (a) $x_j^{(g_{jt})}$; (b) $v_{ij}\left(x_j^{(g_{jt})}\right)$.

regarding each energy service, for the three problem dimensions, i.e., A-Economics, B-Social and C- Environment.

Some of these attributes, are described on Table 3. Then, and based on such attributes, MAVT was used to assist the consumer, by evaluating a set of alternative solutions, obtaining therefore, a set of evaluation tables, each one regarding to an energy service. On Figure 3, it's described an example, regarding a table, correspondent to the energy service "Air Conditioning".

Through the value attributes, obtained by using MAVT, it was used the additive model to aggregate all the evaluation tables. Such model has resulted into 3 different objective functions, regarding each one, to a sustainability's dimension of the problem. The three objective functions considered above, were further optimized, by recurring to an optimization algorithm, based on NSGAII.

As it referred on previous sections, the problem presented here, is from the type of combinatorial nature, whose number of combinations, is related to the dimension of the sample. In this work, it was considered 10 alternatives per energy service.

However, the combinations' number was reduced, since that, the consumer cannot make any choices, given its budget.

Other constraints were also accounted in this work, namely the appliances noise maximal requirements and the air conditioner capacity. Thus, and after being applying MAVT, the problem stated in this work, can be modeled as follows:

$$\begin{aligned} &\max \quad V_D(x), && c/D = A,B,C \\ &s.t\ x \in X && c/V_D(x) = \left[V_A(x), V_B(x), V_C(x)\right]^T \end{aligned} \tag{8}$$

Where x, is the vector of decision, i.e.:

$$x \in X : x \in \left\{ x_{ij}^{(A_{jt})}, x_{ij}^{(\mathrm{B}_{jt})}, x_{ij}^{(C_{jt})} \right\} \wedge \mathrm{t,i},\, j \in \square \tag{9}$$

with,

$$\begin{aligned} &i = \{1,2,..,10\} \wedge j = \{1,2,3,..,7\} \wedge \\ &\wedge t = \left\{ \left\{1,..,n_{A_j}\right\} \cup \left\{1,..,n_{B_j}\right\} \cup \left\{1,..,n_{C_j}\right\} \right\} \wedge n_{A_j}, n_{B_j}, n_{C_j} \in \square \end{aligned} \tag{10}$$

The $V_A(x)$, $V_B(x)$ and $V_C(x)$, defines the objective functions, considering each problem dimension, i.e.:

$$V_g(x) = \sum_{j=1}^{n_j} \sum_{t=1}^{n_{g_j}} v_j(x_j^{(g_{jt})}) \quad w/g = \{A,B,C\} \wedge v_j(x_j^{(g_{jt})}) \wedge n_j, n_{g_j}, \mathrm{t}, j \in \square \tag{11}$$

Therefore, the objective functions are:

$$Economic\ Well-being : \max V_A(x) = \sum_{j=1}^{n_j} \sum_{t=1}^{n_{A_j}} v_j(x_j^{(A_{jt})}) \tag{12}$$

$$Social\ Well-being : \max V_B(x) = \sum_{j=1}^{n_j} \sum_{t=1}^{n_{B_j}} v_j(x_j^{(\mathrm{B}_{jt})}) \tag{13}$$

$$Environment\ Well-being : \max V_B(x) = \sum_{j=1}^{n_j} \sum_{t=1}^{n_{B_j}} v_j(x_j^{(\mathrm{B}_{jt})}) \tag{14}$$

The 1st objective function, were based on previous work [9].

The following approaches (with 2 and 3 objectives) , uses the additive model based on MAVT, to obtain an unique objective function, based on the 2 and 3 objective functions (according to the approach used) pondered by the consumer's concern (ω_g), i.e.:

$$\begin{aligned} V\left(V_A(x), V_B(x)\right) &= \omega_A.V_A(x) + \omega_B.V_B(x) + \omega_C.V_C(x) = \\ &= \sum_{j=1}^{n_j} \left(\omega_A \sum_{t=1}^{n_{A_j}} v_j(x_j^{(A_{jt})}) + \omega_B \sum_{t=1}^{n_{B_j}} v_j(x_j^{(\mathrm{B}_{jt})}) + \omega_C \sum_{t=1}^{n_{C_j}} v_j(x_j^{(C_{jt})}) \right) \end{aligned} \tag{15}$$

The constraints, regarding economic and environment well-being/dimensions, are:

Available Budget ($\eta_{avail.}$):

$$r_1 : \sum_{j=1}^{n_{\dim}} I_j\left(x_j\right) \le \eta_{avail.} \Leftrightarrow \sum_{j=1}^{n_{\dim}} x_j^{(A_{jt})} \le \eta_{avail.} \tag{16}$$

With

$$A_{jt} = \left\{ A_{14}, A_{26}, A_{35}, A_{44}, A_{54}, A_{64}, A_{75} \right\} \wedge n_{\dim}, \mathrm{t}, j \in \square \tag{17}$$

Environment–Noise:

$$r_j : Noise_j \le Max.Noise_j \Leftrightarrow x_j^{(B_{jt})} \le Max.Noise_j \tag{18}$$

With:

$$B_{jt} = \left\{ B_{24}, B_{35}, B_{44}, B_{54}, B_{64}, B_{75} \right\} \wedge n_{\dim}, \mathrm{t}, j \in \square \tag{19}$$

The NSGAII individual framework, presented on Figure 4, regards the approach for the 2 and 3 dimensions of sustainability presented next. For the approach with 2 dimensions, the structure is the same, although considering only 2 dimensions each.

Instead of binary codification, regarding the 1st approach (one objective), NSGAII's codification have used real one. The model, developed here, will be deployed, by using the case study referred before, on previous section.

5. Results & Discussion

5.1. 1st Approach – Single objective model

Based on what was referred before, the 1st approach, has consisted into an EA's single objective approach, whose codification used, was binary, to ensure the existence of a unique individual solution at a time, and regarding appliance's type.

The approach presented here, was implemented by using MATLAB software, given its efficiency when dealing with data, which are organized into matrices. Our previous work ([1]) has allowed us to improve our EA's behavior by testing different objective functions, as well as different parameters values. To assess the quality of EA's solutions, it was used Simplex, which was implemented by using General Algebraic Modeling System (GAMS) software.

The best formulation, achieved at the time from ([1]) , was the one presented on previous section. The results, presented next, were achieved, by using the following parameters:

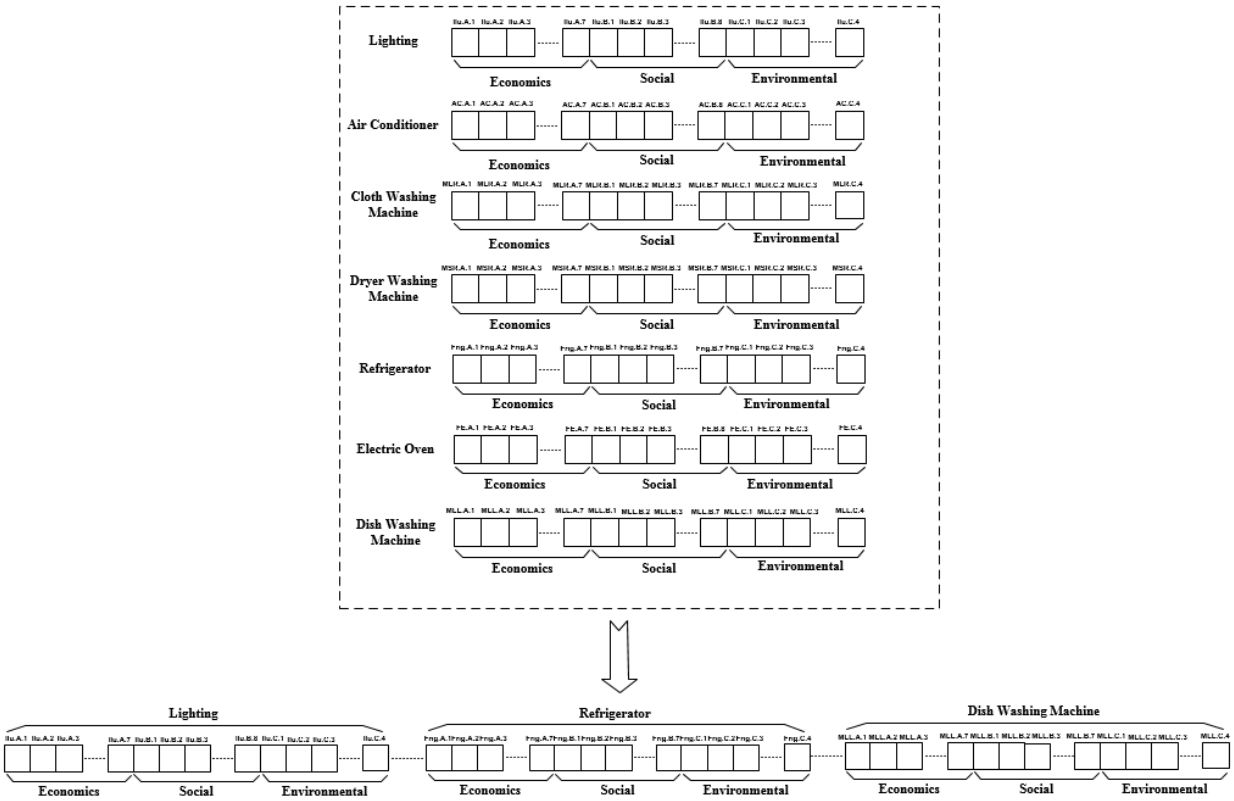


Figure 4: Individual framework (regarding the 2 and 3 objective approachs)

- Population size: 130
- Selection technique: roulette method
- Crossover technique: double point
- Crossover ratio: 0,55
- Mutation technique: bit string
- Mutation ratio: 0,02
- Convergence ratio: 0,002
- Maximum number of iterations/generations: without limits

On previous work ([1]), it was assessed the best fitness function (and objective function), based on the average values of the correspondent $V_R^{(n)}$ objective, by preforming 12 runs/budget.

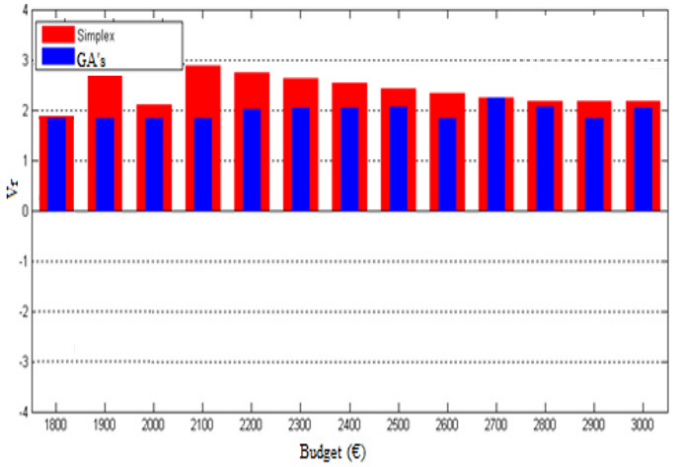


Figure 5: Average values of Vr, considering each value of budget constraint scenario

The EA's behavior, regarding the exploitation of feasible region, was also exploited, by using budget scenario constraint (eq. (7)). Although SIMPLEX, provides the best values, GAs, can also provide good solutions, considering the proximity of both values, By considering the present objective function, EA was also more efficient (on average) than other fitness functions assessed and studied on [1][14]. On Fig.5, it's presented also the average values of Vr (objective/fitness function), regarding the present formulation and considering both methods; GAs and SIMPLEX.

Although SIMPLEX, provides the best values, GAs, can also provide good solutions, considering the proximity of both values, regarding each budget constraint scenario.

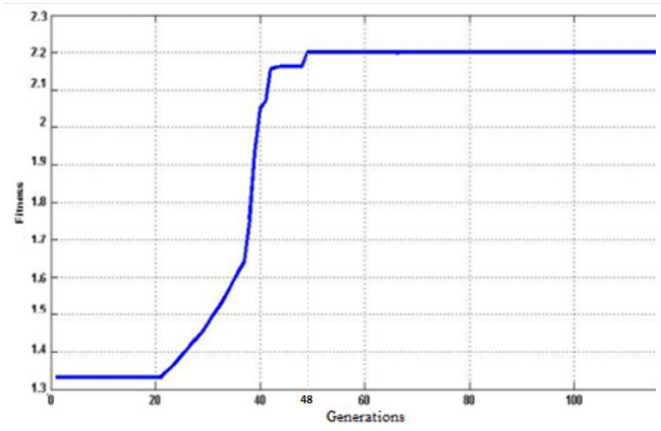


Figure 6: Example of fitness evolution from the best individual/generation (budget scenario of 1900 €)

The performance of GAs is shown on Figure 6, considering a 1900 € budget constraint scenario. On average, the steady state was achieved with 51 generations and considering 12 runs/budget constraint scenario.

On Table 4, it's described a feasible solution achieved with the approach described in this study, regarding a budget scenario constraint of 2700 Euros.

It is also shown the CO_2 avoided emissions (i.e. savings), regarding this solution, by comparing it with the less efficient. Furthermore, the consumer has the information about what brand and model, regarding each appliance type to be acquired, as well as other information, such as the life cycle of each equipement and the savings in terms of electrical energy consumption, by comparying an eficient appliance with a less eficient one.

Table 4. Solution achieved by using EA's approach

Dimension	Stand. Solution Total Invest.	Effic. sol. Total Invest (€)	Invest. Saving (€)	Consum. Saving (€)	CO2 Savings (kg)	Life cycle	Brand	Model
Lighting	15,89	10,55	5,34	58,44	28,50	8000h	Philips	GENIE ESAVER 11W/827
Air Conditioning	368,0	299,0	69,00	1315,60	1315,57	10 yrs	Electro	TC N12KRH PORT
Refrigerator	250,0	529,0	-279,0	704,11	8,50	10 yrs	Indesit	TAN13FFS
Dishwasher Machine	310,0	349,0	-39,00	3,20	6,20	10 yrs	Fagor	1LF-011S
Washing Machine	262,0	294,0	-32,00	6,85	94,80	10 yrs	Zanussi	FLN1009
Oven	170,0	199,0	-29,00	1,30	2,60	10 yrs	Candy	FST100X
Clothes dryer	349,0	419,0	-70,00	11,32	1,800	10 yrs	Indesit	1SL79C
Total:	1724,9	2099,6	-374,7	2100,80	1458,0	-	-	-

Given the stochasticity of the GAs, as a metaheuristic method, it was made a statistical study to evaluate the consistency of the model, which was developed on previous works [1][13].

5.2. 2nd Approach – Multiobjective model

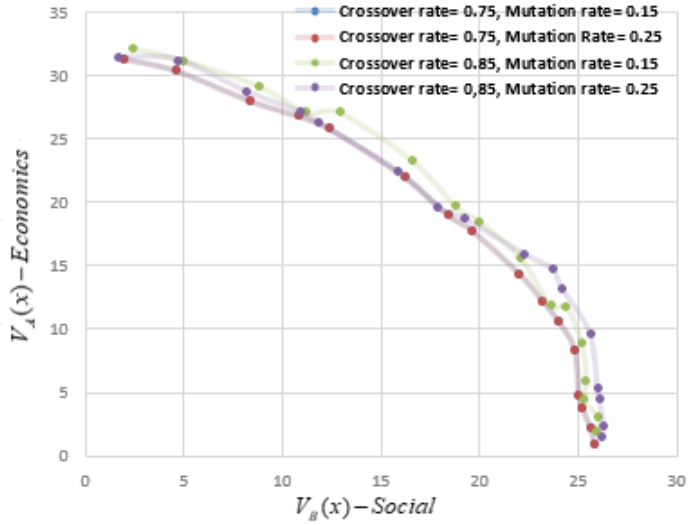
In order to achieve sustainable solutions, and based on Section 4.3, a Multiobjective method, based on NSGAII (Non-Sorting Genetic Algorithm II) ([18]), was coded on Matlab, by accounting the NSGAII's parameters, regarding each NSGAII's phase, namely; selection (roulette), crossover (double point) and mutation (normal random). The other parameters (crossover, mutation rate and initial population) were established after several simulations. Parameters such as the crossover rate, the mutation rate and the population size (120 individuals), were also experimented. With regards to the stopping criteria, it was chosen a parameter value of 80, as being the maximum number of iterations/generations allowed. It was considered several arrangements of crossover and mutation rates of NSGA-II, to obtain a suitable arrangement (Table 5).

Table 5. Values, regarding the arrangements of crossover and mutation rates

Experiment	Crossover Rate	Mutation Rate
1	0.75	0.15
2	0.75	0.25
3	0.85	0.15
4	0.85	0.25

Therefore, it was used the following NSGAII's parameters: population size of 120 individuals, a maximum number of 80 iterations, a mutation rate of 0,15 and a crossover rate of 0,75.

In the Figure 7, it's shown the pareto frontier, for different arrangements of crossover and mutation rates.



After the tuning with NSGAII's parameters, the Pareto frontier in Figure 9, was therefore achieved, with each one of the 16 nodes, representing an optimal solution of the problem. In other words, each node represents a set of individual solutions (appliances from the market), related each one, to an energy service.

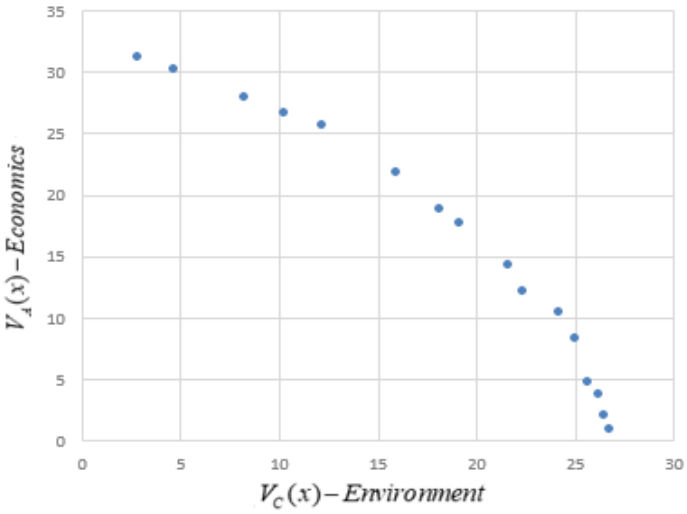


Figure 8. Pareto frontier (last generation) Economical vs Environmental ($\omega_A = 0,73$, $\omega_B = 0,00$, $\omega_C = 0,27$)

Such calculations were performed, by considering a scenario where the consumer only considers the Economic and Environment concerns, i.e., $\omega_A = 0{,}73$, $\omega_B = 0{,}00$ and $\omega_C = 0{,}27$ (Figure 10).

The trade-off, existed between the two dimensions, can also be noted through Figure 8, since the reduction on economic well-being, leads to an increase of the environment well-being.
Table 6, shows the results from one of the nodes achieved here, by assuming a budget constraint value of 2250 €.

Table 6. Solution achieved: Economical vs Environmental
($\omega_A = 0{,}73$, $\omega_B = 0{,}00$, $\omega_C = 0{,}27$)

Dimension	Stand. Solution Total Invest. (€)	Effic. sol. Total Invest (€)	Invest. Saving (€)	Consum. Saving (€)	CO_2 Savings (kg)	H_2O Savings (l)	Brand	Model
Lighting	15.89	09.53	5.34	59.40	28.90	-	GE	EFL23W
Air Conditioning	368.00	299.00	69.00	1320.60	1322.60	-	Whirlpool	PACW9HP
Refrigerator	250.00	529.00	−279.00	708.10	8.70	-	BECKEN	Bc2016 Ix
Dishwasher Machine	310.00	349.00	−39.00	3.20	7.80	429.07	Bosch	SMS25AI00E
Washing Machine	262.00	294.00	−32.00	6.90	94.80	351.02	INDESIT	EWE71252 W
Oven	170.00	199.00	−29.00	1.70	2.33	-	Zanussi	ZZB21601XV
Clothes dryer	349.00	419.00	−70.0	12.30	1.82	-	Electrolux	EDP2074PDW
Total:	1727.89	2098.53	-374.66	2112.20	1466.95	780.09	-	-

Through the values, presented above, we can see that if the consumer, choses the optimal solutions provided by NSGAII, he can save 1737,54 € approximately, having also CO_2 and water savings. Both values are per year, for a life cycle of 10 years.

The same, were performed, by considering a scenario where the consumer only considers the Economic and Social concerns, i.e., $\omega_A = 0{,}73$, $\omega_B = 0{,}27$ and $\omega_C = 0{,}00$ (Figure 9).

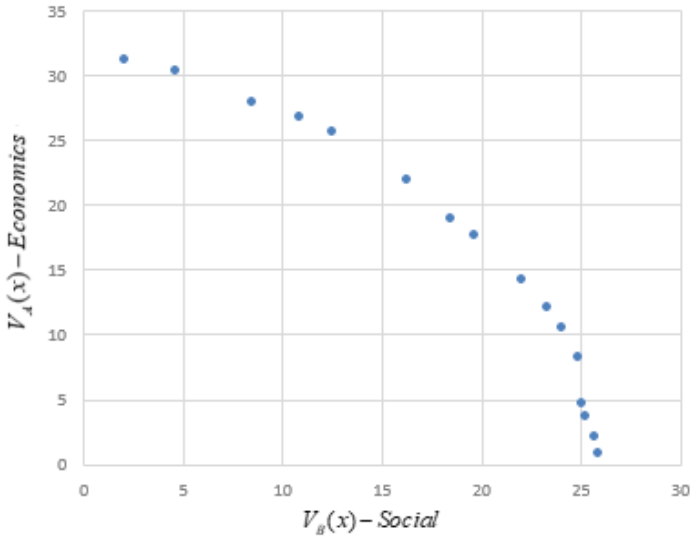


Figure 9. Pareto frontier regarding the last generation; Economical vs Social
($\omega_A = 0{,}73$, $\omega_B = 0{,}27$, $\omega_C = 0{,}00$)

Through the last iteration/generation, it was obtained sixteen feasible solutions, achieving therefore, an example of a Pareto frontier. One of the nodes achieved with this scenario, is shown on Table 7, by considering a budget of 2250 € and a life cycle of 10 years.

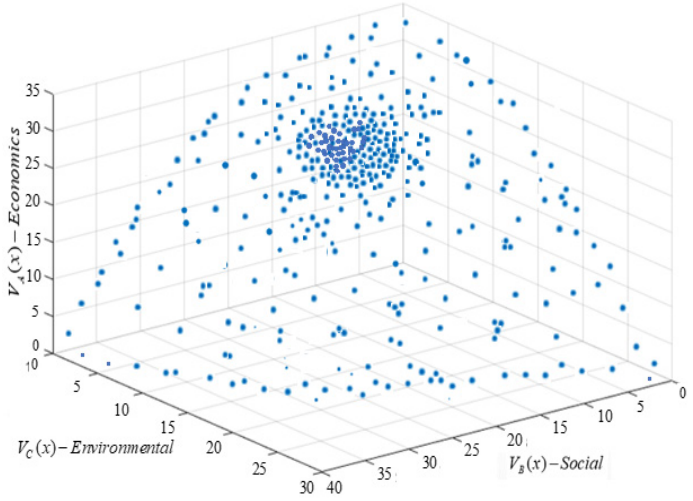
Table 7. Solution achieved: Economical vs Social
($\omega_A = 0{,}73$, $\omega_B = 0{,}27$, $\omega_C = 0{,}00$)

Dimension	Stand. Solution Total Invest. (€)	Effic. sol. Total Invest (€)	Invest. Saving (€)	Consum. Saving (€)	CO_2 Savings (kg)	H_2O Savings (l)	Brand	Model
Lighting	15.89	69.65	5.34	59.40	27.60	-	OSRAM	3316242
Air Conditioning	352.00	279.00	69.00	1319.50	1322.60	-	SAMSUNG	AQV09PSBN
Refrigerator	250.00	529.00	−279.00	708.10	9.72	-	Candy	CFET 6182W
Dishwasher Machine	320.10	351.99	−38.10	2.20	6.90	423.02	BALAY	3VS303IP
Washing Machine	262.00	294.00	−32.00	6.90	95.10	317.04	Siemens	WI12A222ES
Oven	171.00	701.00	−28.30	2.82	2.33	-	Electrolux	EZC2430AOX
Clothes dryer	368.00	449.00	−68.00	10.20	1.82	-	BOSCH	WTE84107EE
Total:	1738.99	2673,64	371.06	2109.12	1458.90	740.06	-	-

According to the values on Table 7, the consumer can save up to 1738.06 €, through the selection of the appliances mentioned above. Additionally, there also savings with CO_2 and water as well, with both values being expressed as savings/years and based on the life cycle considered here.

Based on both trade-offs, presented above, it was considered a scenario with the tree dimensions and their correspondent consumer's relative importance, i.e., Economical vs Social vs Environmental ($\omega_A = 0{,}66$, $\omega_B = 0{,}23$ and $\omega_C = 0{,}11$).

In order to pursue the main goal, referred before, it was performed a scenario where the consumer considers the 3 dimensions of sustainability, i.e., Economic, Social and Environment concerns, i.e.; $\omega_A = 0{,}66$, $\omega_B = 0{,}23$ and $\omega_C = 0{,}11$ (Figure 10).



Through Figure10, it can be seen that the crowding distance, between the final solutions obtained, regarding the obtained Pareto surface, is higher in the region where the Economic dimension is more dominant. Such dominance order is followed by the Social, and at last, the Environmental dimension. Such dominance order, is somehow expected, given the relative importance's values (weight) considered in this case, and regarding each dimension considered, i.e.; $\omega_A = 0,66$, $\omega_B = 0,23$ and $\omega_C = 0,11$.

One of the nodes, represented on that surface, are shown on Table 8, considering a budget of 2250 €, for 10 years of life cycle.

Table 8. Example of a solution obtained from this approach. Economical vs Social vs Environmental ($\omega_A = 0,66$, $\omega_B = 0,23$, $\omega_C = 0,11$)

Dimension	Stand. Solution Total Invest. (€)	Effic. sol. Total Invest (€)	Invest. Saving (€)	Consum. Saving (€)	CO_2 Savings (kg)	H_2O Savings (l)	Brand	Model
Lighting	16.88	49.04	5.35	62.20	28.90	-	Phillips	LEDspotMV
Air Conditioning	352.00	279.00	69.00	1319.50	1315.70	-	SAMSUNG	AQV09PSBN
Refrigerator	234.00	399.00	−265.00	709.30	9.72	-	BECKEN	Bc2016 Ix
Dishwasher Machine	310.00	349.00	−39.00	3.20	6.90	423.00	Bosch	SMS25AI00E
Washing Machine	272.20	249.90	−33.00	5.60	95.10	322.10	INDESIT	EWE71252 W
Oven	171.00	701.00	−28.30	2.82	2.20	-	Electrolux	EZC2430AOX
Clothes dryer	368.00	449.00	−68.00	10.20	1.70	-	BOSCH	WTE84107EE
Total:	1724. 80	2475.94	-262.65	2112.30	1460.22	745.10	-	-

According to Table 8, the consumer, can save up to 1849,65 €, avoiding approximately 1460,22 kg of CO_2 and saving 745,10 liters of water, with both values expressed as savings/years, and based on 10 years as the life cycle considered in this study.

Through the performance of such scenarios, we can see some coherence with the application of ω_A , ω_B and ω_C.

An example is the water savings achieved, where it is more noticed when the correspondent Environment weight (ω_C), is increased.

6. Conclusions & Further Work

This paper follows a research line, where the main goal is to present a method to provide sustainable electrical household appliances from the market to a household consumer. On previous work, and at a first stage, it was formulated a single objective problem, although considering only the environmental impacts (CO_2 savings) and the economic savings, each one, related to the initial investment and energy consumption. Both indicators, were calculated, by preforming the correspondent lifecycle cost assessment (LCCA) of each household appliance, and during the correspondent usage phase.

It was also used Genetic Algorithms (GAs) to achieve several and different feasible solutions, whose quality were tested through the comparison with the results achieved from simplex method.

Then, the approach was formulated by considering a multiobjective problem, although by exploring 2 dimensions at a time, and by recurring to MAVT and NSGA II methods

This approach has used a set of established criteria, in order to perform a pre-selection of candidate solutions existed on market, and suitable to the consumer needs. The purpose was to redefine the decision space, composed by a set of candidate solutions, and based on each type of household appliance considered to be acquired by the consumer.

Furthermore, it was adopted additional criteria, to be further integrated with MAVT, in order to achieve a model with the consumer preferences and based on the 3 problem dimensions presented.

The objective was to improve the consumer well-being, regarding each sustainability's dimension referred above, and based on their importance, given by the consumer.

When achieving the MAVT model, it was also included the ecological impact in terms of water and CO_2 savings, as well as economic issues, such as energy consumption and initial investment savings, based on the lifecycle cost assessment (LCCA) of each household appliance. Social issues were also included, by integrating the consumer's preferences in terms of design, reliability and other issues, including even the visual and thermal comfort. Then NSGAII was applied, by achieving sustainable solutions from the market, that maximizes the tree dimensions referred before.

The approach presented here, also allows a consumer, to achieve a set of savings regarding issues such as; CO_2 emissions, energy and water consumption.

As further work, this approach could be applied into other energy services/household appliances, for instance, with a relevant impact in terms of sustainable development (e.g. information technology's equipment's such as computers, printers, among others).

References

[1] R. Santos, J. C. O. Matias and A. Abreu, "An approach to provide efficient choices to a household consumer," 2019 International Young Engineers Forum (YEF-ECE), Costa da Caparica, Portugal, 2019, pp. 46-51. doi: 10.1109/YEF-ECE.2019.8740824

[2] IEA: Energy Efficiency 2017 – Market Reports Series, OECD/IEA (2017)

[3] Gul, M., Patidar. S. (2015). Understanding the energy consumption and occupancy of a multi-purpose academic building, In Energy and Buildings, Volume 87, Pages 155-165, ISSN 0378-7788

[4] ADENE: Manual da Etiqueta Energética, ADENE, Lisboa ISBN: 978-972-8646-36-3 (2017)

[5] Wong, I., L., Krüger, E., Comparing energy efficiency labelling systems in the EU and Brazil: Implications, challenges, barriers and opportunities, In Energy Policy, vol.109, pp. 310-323, Springer, Heidelberg (2017)

[6] Fell, M.: Energy services: A conceptual review, Energy Research & Social Science, vol.27, pp.129--140, Springer (2017)

[7] Hoxha, E., Jusselme, T. (2017), On the necessity of improving the environmental impacts of furniture and appliances in net-zero energy buildings, In Science of The Total Environment, Volumes 596–597, Pages 405-416, ISSN 0048-9697

[8] Ting, T.O.; Rao, M.V.; Loo, K.C. (2006) A novel approach for unit commitment problem via an effective hybrid particle swarm optimization. IEEE Trans. Power Syst., 21, 411–418.

[9] Ferreira, F.A.; Lemos, F.A.B. (2010) Unbalanced electrical distribution network reconfiguration using simulated annealing. In Proceedings of the 2010 IEEE/PES Transmission and Distribution Conference and Exposition: Latin America (T&D-LA), Sao Paulo, Brazil, 8–10; pp. 732–737.

[10] Agrawal, P.; Rao, S. (2014), Energy-aware scheduling of distributed systems. IEEE Trans. Autom. Sci. Eng. , 11, 1163–1175.

[11] Ko, M.J.; Kim, Y.S.; Chung, M.H.; Jeon, H.C.(2015), Multi-objective design for a hybrid energy system using genetic algorithm. Energies, 8, 2924–2949.
[12] Randall, M., Rawlins, T., Lewis, A., Kipouros, T., Performance Comparison of Evolutionary Algorithms for Airfoil Design, In Procedia Computer Science, vol. 51, pp. 2267-2276, Springer Heidelberg (2015)
[13] R.S. Santos, J.C.O. Matias, A. Abreu, F. Reis, Evolutionary algorithms on reducing energy consumption in buildings: An approach to provide smart and efficiency choices, considering the rebound effect, Computers & Industrial Engineering, Volume, 126,2018,Pages 729-755,ISSN 0360 8352, https://doi.org/10.1016/j.cie.2018.09.050
[14] Santos, R., Abreu, A., Matias, J.C.O.,Energy Efficiency in buildings by using evolutionary algorithms: An approach to provide efficiency choices to the consumer, considering the rebound effect, IFIP International Federation for Information Processing 2018, 2018-03-29, book-chapter Part of ISBN: 978-3-319-78574-5
[15] Pombo, O., Allacker, K., Rivela, B., Neila, J., (2016) Sustainability assessment of energy saving measures: a multi-criteria approach for residential buildings retrofittingda case study of the Spanish housing stock, Energy Build. 116 pp.384e394.
[16] Aghay Kaboli , S. Hr., Selvaraj, J., Rahim, N.A.,(2017) Rain-fall optimization algorithm: A population based algorithm for solving constrained optimization problems, Journal of Computational Science, Volume 19, Pages 31-42, ISSN 1877-7503, https://doi.org/10.1016/j.jocs.2016.12.010.
[17] Cortés, P., Muñuzuri, J., Berrocal-de-O, M., Domínguez, I., (2018) Genetic algorithms to optimize the operating costs of electricity and heating networks in buildings considering distributed energy generation and storage, Computers & Operations Research, Volume 96, Pages 157-172, ISSN 0305-0548, https://doi.org/10.1016/j.cor.2018.01.020.
[18] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," in IEEE Transactions on Evolutionary Computation, vol. 6, no. 2, pp. 182-197, April 2002. doi: 10.1109/4235.996017
[19] Chuah, J. W., Raghunathan, A., Jha, N. K. (2013), ROBESim: A retrofitoriented building energy simulator based on EnergyPlus, Energy Build., vol. 66, pp. 88–103.
[20] Asadi, E., Silva, M. G., Antunes, C. H., Dias, L. (2012), Multi-objective optimization for building retrofit strategies: A model and an application, Energy Build., vol. 44, pp. 81–87
[21] Mauro, G.M., Hamdy, M., Vanoli, G.P., Bianco, N., Hensen, J.L.M. (2015) A new methodology for investigating the cost-optimality of energy retrofitting a building category, Energy Build. 107 pp. 456 e 478.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 1, 408-417 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Resource Selection Service Based on Neural Network in Fog Environment

Nour Mostafa*

College of Engineering and Technology, American University of the Middle East (AUM), Egaila, Kuwait

ARTICLE INFO

Article history:
Received: 14 January, 2020
Accepted: 18 February, 2020
Online: 25 February, 2020

Keywords:
Cloud
Fog
Fog-to-Cloud
Fog-to-Fog
IoT
e2e delay

ABSTRACT

As an emergent technology in Internet of Things (IoT), the ultimate target of fog computing is to provide a widely distributed computational resources and data repository closer to the network edge providing heterogeneous systems both in terms of software and hardware. The fog system must have the capability to deal with huge number of resources and users at the same time, such increased size sometimes presents the issue of performance degradation. Therefore, the fog should be able to support adaptability, scalability, and extensibility to avoid such degradation by adopting efficiently and effectively an optimal resource selection and allocation model. As many fog users have limited interest in, or knowledge of fog middleware issues, it follows that in a large fog environment the best approach would be to have an automated system for resource selection and allocation. Such an approach eliminates the need for user intervention. This paper proposes a fog resource selection service based on neural network to perform resource selection tasks by coordinating with the metascheduler. Five different selection algorithm were used to evaluate the prediction model for resource selection. In addition to introducing a history update and management algorithm to manage and control the storage of the history log records.

1. Introduction

Fog and IoT systems are handling huge numbers of users, resources and tasks. The fog and IoT environment are widely distributed with great diversity, therefore, it is expected that resource selection will be carried out by run time estimates. With the increase size of the fog and IoT and also with its wider acceptance by the 'ordinary' users, some form of automatic resource selection system is desirable. This resource selection system should have the capability to evaluate available resources on the basis of specific criteria as determined by the user.

Because of its huge volume, the fog requires special services for its efficient utilization. With the expansion in the size of the fog and IoT, users are facing new challenges of resource selection to meet their computational requirements without getting into the fine details of available resources. Although, evolving very rapidly and becoming widely accepted, normal users of the fog and IoT find it very difficult to use middleware technologies. Middleware services provide a means to virtualize and aggregate resources. Fog resources are heterogeneous and very complex and require intelligent solutions for resource discovery and characterization. Therefore, in fog environment, the network and fog designers are facing a serious challenge to grant efficient utilization of such huge and widely distributed resources. Such increased size sometimes presents the issue of performance degradation. Therefore the fog system should be able to support adaptability, scalability, and extensibility to avoid such degradation. In order for the fog system to scale well with the increase size of resources and users, we propose a Fog Resources Selection Service (FResS) for fog systems based on neural network that will meet a user's preferences in an optimal response time. The proposed strategy presents a predictive resources selection model to predict and determine the optimal resource to fulfill the user's task. The work introduced in [1] has been extended in this paper where the IoT framework is integrated within the FResS, in addition to testing and analyzing different selection algorithms. Moreover, a history update and management algorithm has been introduced to manage and control the storage of the history log records. This situation requires that the fog user should be provided with a helper service to facilitate resource selection. Resource selection can be done by taking into consideration either, cost, time, cost and time, performance, or any other constraint. This resource selection service can have its own resource selection algorithms and can also be deployed at varying distribution levels. Currently the majority

*Corresponding Author: Nour Moustafa, nour.moustafa@aum.edu.kw

of metaschedulers are deployed at the cloud level which will be a bottleneck for larger fog and IoT systems. A higher distribution level will alleviate this bottleneck. The future fog and IoT is also intended to be more user focused.

The advanced fog selection for task execution is enabled by the resource selector which uses the FResS to make a rum time prediction to select the best host/resource. Due to the resource-limited fogs, the proposed model avoid incurring overhead when executing the task. In addition to decreasing the overall end-to-end (e2e) latency of the system. The proposed FResS model select resources based on criteria mentioned by the user, which provide QoS levels that meet the users' expectations. Experimental performance evaluation shows an overall improvement in terms of bandwidth consumption and task execution times, in addition to maintaining adequate QoS levels.

The reminder of the paper is organized as follows: Section 2 presents background reading and related work. To provide a deeper understanding of this area, section 3 introduces the fog prediction and selection scheme. Section 4 presents the multi-level load balancing technique. Section 5 presents the evaluation results using workloads and simulations respectively. Finally, the paper concludes with section 6, focusing on the contribution along with future recommendation.

2. Related Work

The emergence of cloud, fog and IoT approaches as a comparatively new approaches to distributed computing, which has become popular in the last few years. These approaches addresses the problem of organizing large scale computer societies, called Virtual Machine, which are able to use and share large sets of distributed resources (e.g. computational resources and data repositories) among them. With the emergence of several cloud and fog service providers, it is expected that system performance and run time estimates will be used for resource selection, which make organization to move from one service provider to another to meet their needs. Predictions are a very important step towards automatic resource management [1][2]. As the capabilities of IoT devices are limited, more attention is required when allocating tasks on these devices to assign the workload optimally on available resources.

In [3] authors aims to enable parallel execution by splitting the allocated services on the available resources. To achieve low delay on service allocation the proposed model distributed/matched the desired services among the available Fog-to-Cloud (F2C) resources. The author in [4] proposed a resource allocation strategy for resource selection by adopting the Price Timed Petri Nets (PTPNs) which allow the user to select the resources autonomously based on the price cost and time cost to complete a task. In addition to sorting the resources into groups based on their credibility, then the users will be assigned to different groups because the credibility that they valuate the resources is different. The authors in [5] proposed an online algorithm for task allocation in mobile IoT networks by considering two scenarios to address the scenario of tasks information is unknown and incomplete local task information, respectively. The task allocation decision is done based on the current network status, energy consumption in order to satisfy the mobile device battery capacity constraint and to minimize the energy consumption of the current location.

In [6], the authors proposed an online task scheduling scheme called FairTS by adopting and combining the deep reinforcement learning (DRL) and dominant resource fairness (DRF) techniques. The proposed scheme views the agent's policy, take the current state using neural network to learn from experience, and outputs the action selection probability vector. The agent observe the resource availability, pick a task, and decide its fog resource allocation based on a resource availability that meet the task completion deadline to achieve resource fairness among tasks.

The author in [7] proposed offloading scheme, where the task data submitted by the user is uploaded to its nearby fog node, then the fog node will take a decision for offloading the task to its neighbor fog node or remote cloud to satisfy the delay deadline received from the end user. In [8] the authors used the machine learning as a prediction model to predict cloud processing resource consumption. The proposed model uses an artificial intelligence algorithm for future request prediction by dividing the incoming request into multiple consumption classes depending on their consumption amount, then the request is distributed among processing unites. For an incoming request, the adopted algorithm is used to predict the resource amount consumed to complete the task, a placement algorithm is used to allocate it to the best processing unit.

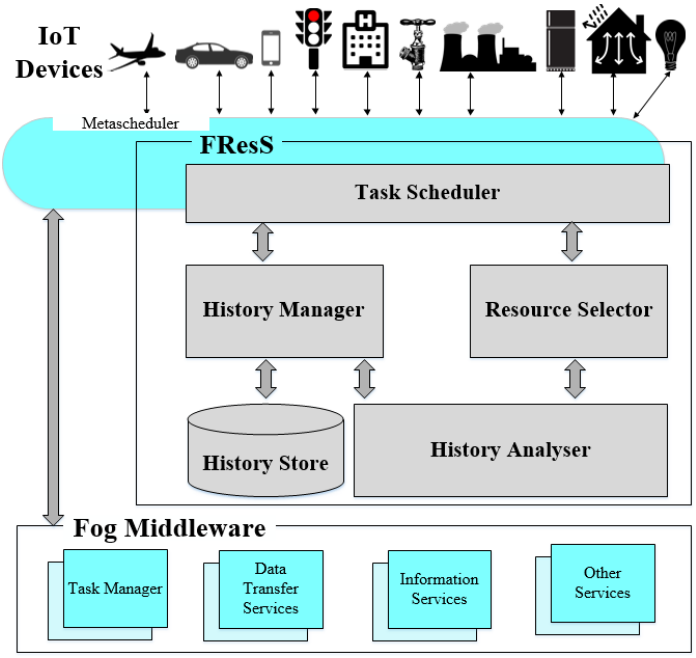
Although there exists research investigating the resources selection and allocation issues in fog computing, however it is still premature. As mentioned earlier, this project resulted in the development of two novel components: a run time prediction model and a distributed architecture of a Fog Resource Selection Service (FResS), which provides a degree of autonomy in determining the optimal resources to execute a particular task.

3. Fog Prediction and Selection Scheme

In our paper [1] a novel approach has been introduced in fog computing, namely, a Fog to Fog (F2F) resource selection algorithm called FResS, that provide automatic resource selection system. This resource selection system have the capability to evaluate available resource on the basis of specific criteria as determined by the user. The idea was to develop a new technique which require no user input. This work resulted in a run time prediction model based on historical execution logs and a fog resource selection service. The main modules of the FResS framework are: Task Scheduler, Resource Selector, and History Analyzer. These modules perform different resource management functions, the most obvious of which is resource prediction and selection. All of these modules are discussed separately in the coming sections.

A detailed architectural diagram of FResS is shown in Figure 1 highlighting the fact that FResS works along with the metascheduler. The FResS architecture was developed by keeping in mind its distribution level, which can be either IoT level, or fog level. The recommended distribution level is the IoT level, if deployed at the IoT level, FResS stores historical data related to the single IoT, it makes the management of execution logs straightforward because the overall volume of data will be very small. It will also provide fast and accurate predictions, again

because of smaller and related data. If deployed at the fog level it can become a single point of failure and a performance bottleneck. Implementation details will be very much similar for the different levels of distribution.



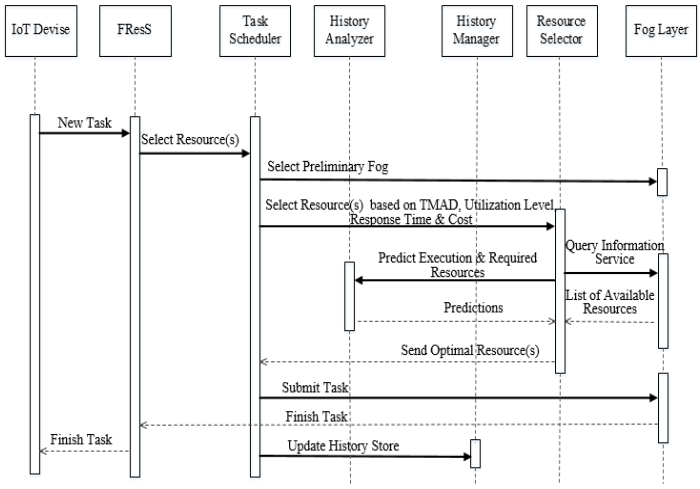
architecture

3.1. Task Scheduler Module

This is the main component interacting with the fog middleware through a metascheduler. Its implementation depends on the underlying fog middleware. Depending on its distribution level, whether it is IoT level or fog level, the Task Scheduler Module can have different implementation details. At a minimum it provides interfaces between the metascheduler and the FResS components, but it can also act as a metascheduler in itself. For evaluation purposes a simplified Task Scheduler Module was implemented by removing all external constraints.

The Task Scheduler sits between an IoT devise/user and the fog middleware. It provides interfaces to the IoT, the fog middleware, and other modules of the metascheduler. The sequence diagram shown in Figure 2 explains the sequence of events during a typical task submission process. The Task Scheduler Module and the metascheduler module coordinate with the lower layers of the fog to execute different tasks. Lower layers of the fog provide basic services i.e. data transfer, task submission, task management, and information services. Once a task is submitted by the user to the Task Scheduler, it will be connected to a preliminary fog based on users' location, and then the Task Scheduler Module forwards it to the Resource Selector Module.

The configuration plays a key role in selecting resources for each task and predicting its execution time. If the task selector returns more than one resource for a particular task, then an optimal resource will be selected by the Task Scheduler to fulfill users' request based on utilization level, which allow the users' preference to be considered while assigning the task to resource(s) which is more suitable for large scale domain such as fog computing. The progress of the tasks and resources will be constantly monitored by the proposed model, which take decisions based on that. The History Manager will receive the execution logs from the Task Scheduler after every execution, and subsequently predictions will be made for future tasks using these logs. The sequence diagram in Figure 2 shows the interaction among the different modules of FResS during a typical task submission process.



FResS model

3.2. Resource Selector Module

The Resource Selector Module analyses the workload of the incoming task to provide automatic resource selection by using the run time predictions. The Resource Selector depending on the configuration, along with their run time predictions, will forward a list of selected resources to the Task Scheduler. The Task Scheduler Module returns the resource list to the metascheduler to complete task submission. The resource selection on the fog is a multi-step process. The first step tests the eligibility of the discovered resources against the essential QoS criteria, e.g. privacy, memory, cost, availability, and delay. The list of the selected resources should be capable of executing the current task if there is no deadline constraint. Step one is performed by the metascheduler on its own.

In step two, the ability of resources to meet deadlines will be evaluated by evaluating run time predictions. For optimum task allocation, these predictions are then utilized by the metascheduler. To make optimum resource selections the metascheduler also considers additional constraints of user preferences, data locality, co-allocation, cost, workflow constraints, and advance reservation. These additional constraints are also evaluated by the metascheduler after receiving a ranked list of resources along with predictions from FResS. Predictions help the metascheduler to make a decision on the required duration of advance reservation. The co-allocation decision is facilitated with this predicted knowledge of expected run times.

The Resource Selector Module receives its input from the Task Scheduler Module along with the currently available resource list. It forwards incoming tasks to the History Analyzer Module to generate run time predictions for the available resources. The History Analyzer Module then uses algorithm 6, discussed in the next section, to generate run time predictions. For the evaluation

of FResS an experimental strategy was devised in which five algorithms 1, 2, 3, 4, 5 were used by the Resource Selector Module for resource selection. These algorithms are named as history, static, random, history plus static, and history plus random.

Data: workload history
Input: available resource list
Input: task submission description file
Result: selected resource
initialization;
getPredictions(new task submission description file, available resources);
selected resource is one with the minimum run time prediction;
return selected resource;

Algorithm 1: Selecting resource using history

Data: workload history
Input: available resource list
Input: task submission description file
Result: selected resource
initialization;
getPredictions(new task submission description file, available resources);
selected resource is one with the highest mips rating;
return selected resource;

Algorithm 2: Selecting resource using static performance capability

Data: workload history
Input: available resource list
Input: task submission description file
Result: selected resource
initialization;
getPredictions(new task submission description file, available resources);
select a resource randomly from the available resource list;
return selected resource;

Algorithm 3: Random resource selection

Data: workload history
Input: available resource list
Input: task submission description file
Result: selected resource
initialization;
getPredictions(new task submission description file, available resources);
select 10 resources with the least run time predictions;
from these resources pick resource with highest mips rating;
return selected resource;

Algorithm 4: Selecting resource using history and mips rating

Data: workload history
Input: available resource list
Input: task submission description file
Result: selected resource
initialization;
getPredictions(new task submission description file, available resources);
select 10 resources with the least run time predictions;
from these resources pick any resource randomly;
return selected resource;

Algorithm 5: Selecting resource using history and random selection

Selecting the fastest resource depending on its static number crunching capacity is a straightforward solution to the resource selection problem. It is approximately comparable to the Condor matchmaking technique [9]. It is mentioned above as static resource selection, which suffers from the problem that it does not have a feedback mechanism to inform the system about any anomaly. The history based resource selection on the other hand has a feedback mechanism to rectify any anomaly resulting from changed resource capability.

The random resource selection results in an even distribution of tasks among candidate resources without considering the difference in their capabilities. On heterogeneous fog system, it is not a desired outcome where high power resources are required to get more tasks. The random resource selection also does not provide any estimates of expected execution time, leaving the user in a 'wait and see' situation. The historical resource selection is combined with last two algorithms, 4 and 5 to overcome the drawback of static and random resource selection. Another shortcoming of the historical resource selection is the uneven load distribution similar to the static resource selection. However, the presence of a feedback loop shown in Figure 4 reduces the severity of this problem. The performance of all these algorithms will be evaluated in detail in the simulation result section. Once similar tasks are found, run time predictions are generated for all resources in the list. The finished list is returned to the Resource Selector Module.

3.3. History Analyzer Module

The incoming task along with the list of available resources is forwarded to the History Analyzer Module to generate run time predictions by examining through the history execution logs to find similar tasks using algorithm 6, this process is conducted using Artificial Neural Networks (ANNs) [10]. Once similar tasks are found, run time predictions are generated for all resources in the list. The complete list is then returned to the Resource Selection Module. ANNs are suitable for training over hundreds or even thousands of passes of data sets[11]. The increased size of the fog and IoT systems brings new challenges of scalability and extensibility and the need for automated processing becomes clear therefore, accurate results can be achieved. The design of ANNs generally comprises three layers, input, middle (or "hidden"), and output layers [10]. The input layer is the starting point where the data enters the system. The input data are passed to the hidden layer for processing which is the intermediate processing unit, and then passes the new signal onto the output layer. The learning phase plays a key role in neural network for processing information

using the interconnection weights. The strength of the input data is measured by weights, therefore, the process of finding the best set of weights for the neural network is referred to as training or learning. In order to get accurate results, the weights are modified at the input layer by passing input values, and then the network's predictions will be measured to check how the predictions' values are close to the training sets. In addition, the accuracy of the prediction results is compared to the actual system results as shown in the simulation result section 5, Figure 3 shows a single unit network.

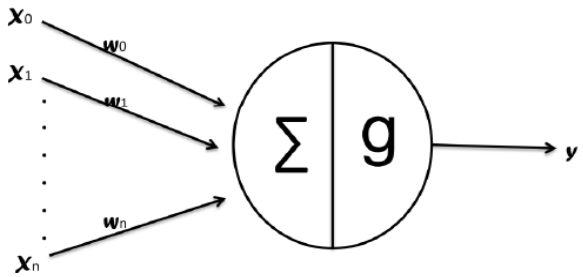


Figure 3: An Artificial Neuron Unit [12]

An amount proportional will be changed to the difference between the desired output and the actual output to achieve weight adjustment [13]. The general mathematic definition [14] is as shown in (1) describes the relationship between the input and output mapping:

$$y(x) = g(\sum_{i=0}^{n}(w_i x_i)) \quad (1)$$

the input of each neuron is represented by x with (n+1) input dendrites (x_0….. x_n) and the axon $y(x)$ computes the output. In order to supply a node activation the weighted signals are initialized with (w_0….. w_n). The activation function is the identity function and is represented by g which computes the weighted sum of the inputs from other neurons and outputs it. The aim is to be able to produce the output within desired accuracy matching the input pattern [15].

After the historical execution logs are examined by the History Analyzer to generate estimates for the required resources by finding similar task with sufficient accuracy, once predictions are generated, the complete list will be forwarded to the Resource Selector to determine the optimal resources to execute the task. Algorithm 6 forms the core of FResS, which is used to generate rum time prediction for new tasks on the available resources.

The process of identifying the resource id dependent on five parameters: delay, price, previous execution time, memory requested, and submission time. Workloads can be modelled at any level but it is found that modelling at a user level is more realistic, this is because users of fog networks tend to repeat same tasks using the same data [12] [16]. Therefore, the results are interesting since prediction of required resources of the new task will be made using past execution parameters.

The proposed algorithm combines the static and random algorithms with the historical resource selection to overcome their disadvantages, e.g. in certain situations there is a possibility of selecting the same resource repeatedly by the Task Scheduler, creating an uneven load distribution. Consequently, the presence of the feedback loop as shown in Figure 4 overcome the severity of this problem.

Data: user preferences and task log history
Input: available resource list
Input: new task submission description file
Result: run time prediction for the resource list
Start
While Taskset <> empty
do
  process request for Task(s) k_i;
  if (Ki submitted by new user/device) then;
    Resource Selector send available resources to Task Scheduler;
    Task Scheduler send user preferences to Task Manager;
    if (Task Manager found optimal resource(s));
      send back optimal resource(s) to Task Scheduler;
      execute k_i on the optimal resource;
  else if (Ki submitted by existing user/devise) then;
    NN predict required resources for k_i;
    Send back predicted resource(s) to execute k_i;
    execute k_i on the predicted resource;
  else
    execute k_i on the preliminary fog;
  update Task Manager log history;
  select next task;
  end
end

Algorithm 6: Making predictions for run time

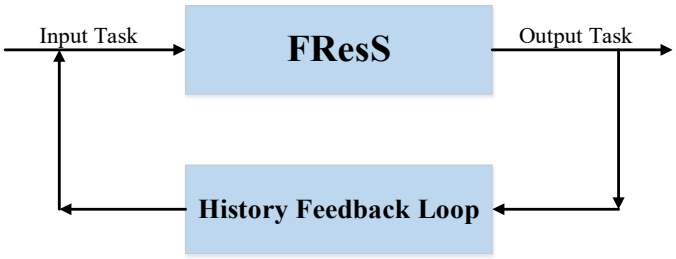


Figure 4: History as a feedback loop in the scheduling process

On the other hand the history management algorithm (algorithm 7) works separately and is responsible for keeping the history up to date. It means that it stores new history records after successful executions and removes older history records which is updated continuously after every execution and subsequently used by ANN to make predictions for future tasks. Subsequently, the increasing number of the historical executions records will increase the accuracy of predictions [17][18].

3.4. History Manager Module

The History Manager is interfaced with the Task Scheduler and history database, it is responsible for keeping the history data up to date. History updates are performed by using algorithm 7. The History Manager Module controls the storage and removal of

execution records as required. Algorithm 7 performs history management by storing completed tasks data in the database as soon as they are received. After this step, a query is sent to the database to retrieve the size of the node related to the new task, just stored. If the size of this node is less than, or equal to, the maximum history limit then no action is taken. Otherwise, if the size of this node is larger than the maximum, then the oldest member of this node is deleted. This module gathers different QoS data e.g. delay, cost, previous execution time and memory requested etc.

Data: workload history
Input: execution log of completed job
Input: task submission description file
Result: history updated
initialization;
store record to the history database;
if size of this cluster is more then the size limit then
  delete the oldest record;
end

Algorithm 7: Update history

3.5. Complexity of Algorithms

Algorithms 6 and 7, presented above, form the core of FResS. The predictions algorithm (algorithm 6) is used to generate run time predictions for new tasks on the available resources. On the other hand the history management algorithm (algorithm 7) works separately and is responsible for keeping the history up to date. It means that it stores new history record after successful executions and removes older history records.

The complexity of predictions for algorithm 6 is of $O(n)$ where n is the size of available resources. On the other hand, complexity for history management algorithm 7 is of the $O(1)$. These algorithms will become part of the task submission service. The history management algorithm 7 can be executed offline, after the task is finished, so it's speed is not very critical to the overall execution time because it is not contributing to the overall delay. Algorithm 6, on the other hand, will contribute to the total delay in the execution time.

3.6. Fog Workflows and FResS

Predictions are utilized by the task management system to select resources to execute incoming workflows consisting of many tasks. These workflows can be very simple, consisting of only one task, as shown in Figure 5, or can be very complex, consisting of many interdependent tasks, as shown in Figure 6. These tasks can have dependencies which restrict them to execute in a certain sequential order or execute concurrently. Predictions for each task within a workflow are generated separately.

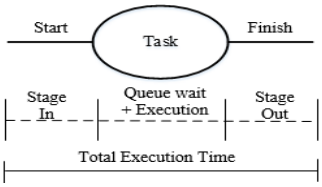
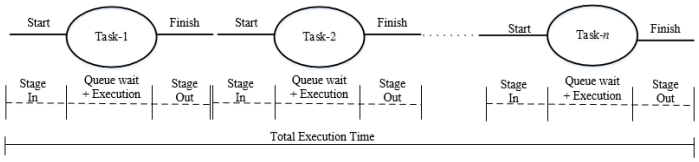


Figure 5: Workflow consisting of only single task



Quality of Service (QoS) refers to the level of performance and service that a given user will experience at a certain point in time, when starting a certain operation on a certain instance of a service. QoS support refers to the possibility of a certain level of required performance being available from a certain resource or not. QoS is very important to fog applications which are run collaboratively in real time. QoS requirements can be mentioned in terms of cost, delay, scalability, fault tolerance, etc.

Normally QoS requirements for these workflows are applied to the entire workflow. To achieve the best timing performance it is recommended that the workflow is partitioned in such a way that parallel activities finish at the same time. Because of the global nature of the fog, locality will also play its role to achieve best QoS results. Sometimes if the user wants to see intermediate results, then QoS can also be mentioned at the task level. Achieving QoS either at task level or workflow level boils down to the task or activity level. To achieve QoS at the workflow level requires that tasks are managed properly. Incorrect management of these tasks can damage QoS at a very high level.

The proposed FResS model is responsible for the overall execution of workflows. These executions, in certain situations, may also require advance reservation and co-allocation. The FResS performs its task comparatively easily when accurate run time predictions are available for tasks within a certain workload. Advance reservation is a technique in which resources are reserved for a certain task to start at a certain time and execute for certain duration. On the fog, advance reservation will be expensive so its accuracy is of prime importance. With the predictions already available, the FResS can reserve resources for accurate time duration, resulting in an efficient usage of resources.

4. A Multi-Level Load Balancing technique in Fog Computing

Internet traffic growing exponentially and congested with thousands and millions of devices [20], quality services should be provided with higher availability by the service provider. With the increase number of services in fog and IoT computing the network load is expected to increase which will create a new challenge to the researchers and network designer. The aim is to achieve better performance in terms of the total time to execute the tasks load balancing to accomplish better QoS, in addition to increasing the number of task acceptance. The fog computing attempts to integrate multiple distributed and heterogeneous resources, which are normally under separate administrative domains resulting in under-utilization or over-utilization. A multi-level load balancing approach has been proposed to provide better load balance. As mentioned above in algorithm 6 user preferences will be considered for the incoming task. Therefore, satisfying the below metrics and parameters optimally will led to achieve better performance of the fog system.

4.1. Metrics

- Cost: IoT and fog users are interested in the best performance at the lowest cost, but on the other hand the resource owners are interested in the overall system throughput. Therefore, the efficient algorithm should consider the cost parameters by scheduling the incoming task to a resource based on the cost.
- Response time: is the total execution time taken by the task to be completed, it starts from submitting the task by the user or device and end with receiving the completed task from the service provider. A smaller response time is always desirable.
- Delay: Researchers and service providers stress on the importance of guaranteeing acceptable delay, since this is an important metric in terms of QoS, the proposed model adopted the tenant maximum acceptable delay (TMAD) [21]. TMAD sorts the incoming tasks to a higher priority task or lower priority by considering different prioritization of users based on their important for the system according to the subscription plan each user pay for.
- Scalability: bottlenecks in fog computing system arises from the huge number of users, resources, and services, therefore to be able to support scalability, and extensibility, this is an important metric which determine if the system is able to achieve better load balancing with a restricted number of resources.
- Fault Tolerant: the larger the system, the more frequent failures result, IoT and fog systems should provide the ability to perform correctly in case of failure by having a backup

5. Simulation Results

The performance of the proposed solution is evaluated in this section by carrying out some experiments using CloudSim [21], selected for its flexibility and wider acceptance by the cloud and fog community. The proposed model performance was evaluated by comparing it with existing scheme and highlighting the differences and advancement. IoT devices were simulated as geographically distributed nodes. The fog network represented as a graph and storage sites were modeled as a set of nodes. A computing power, storage capacity, narrow and broad bandwidth, and memory have been configured to all nodes to be close to reality. The number of tasks, and capacity of storage nodes were varied to show the complexity of the proposed model and simulate different scenarios. The simulator defined 2 cloud sites, 5 fog sites, and 150 IoT devices, 5000 task requests, and 500 to 2000 mb/sec the connectivity of the bandwidth.

Scheduling with predictions helps the scheduler to select the best resource(s) for a certain task and it can be compared with two other techniques i.e. random and static. In a random selection the scheduler selects any resource from the qualifying resources. Qualifying resources are those which meet the minimum requirement criteria for a certain task. On the other hand static selection is one in which the resource with the highest number crunching speed is picked. In a real world scenario, number crunching speed can be based on any popular benchmark. Experimentation was conducted with CloudSim, which defines resource capability in term of Million Instructions Per Second (MIPS), hence in the static resource selection, resource with the highest MIPS rating will be selected from the qualifying resources.

In a deadline based scheduling scheme the function of the scheduler is to meet a deadline specified by the user. It is quite possible that more than one resource can meet this deadline, established by comparing the deadline with the predicted time. In such cases the scheduler can select the fastest resource, which is referred to as history based selection. It is expected that history based resource selection will create an uneven load distribution on the fog system. There are two other options: (1) select the fastest resource from the selected fastest resources, (2) select any resource randomly from the qualifying list. These options were given the name of history plus static and history plus random.

As mentioned in section 3.2, it can be seen that in all, there are five scheduling algorithms, i.e. History, Random, Static, History plus Random, and History plus Static. These algorithms were evaluated using simulation. There are two aspects of the performance of the scheduling algorithms: (1) the execution of tasks within a minimum possible time, (2) the distribution of loads evenly among the available resources. The first experiment was designed to compare the performance of these algorithms in terms of the total execution time to complete the same tasks. In this case the same tasks were submitted to the fog by using each of the five algorithms, one by one, and the total time to complete these tasks was recorded. This experiment was repeated for the tasks created by all workloads, one by one. Results were plotted in the form of bar graphs for each workload based tasks, separately, as shown in Figure 7.

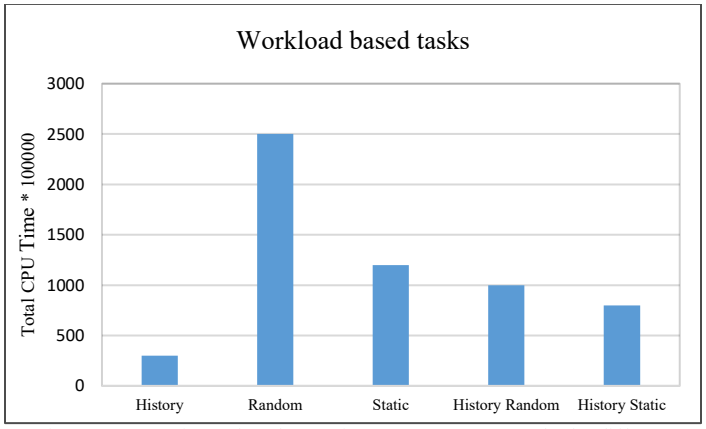


Figure.7 Comparison of total CPU time to execute 5000 tasks for different resource selection algorithms

Different scheduling algorithms were used to evaluate the performance of the predictor i.e. history, static, random, history plus static and history plus random. It was seen that overall execution time was minimum for the static and history based scheduler but it also created issues of load imbalance. History plus random and history plus static provided better load balance, but took larger overall time to finish the same tasks. It was concluded that selective task scheduling will create a load imbalance but a history based solution will operate to fix this imbalance because of the presence of a feedback loop. Random and static schedulers are missing this feedback loop, and hence lack the self-healing property of history based algorithms. Overall performance of FResS was found to be better when using the historical technique, or any hybrid technique that incorporated the historical technique

The accuracy of predictions was also compared with user provided run time estimates. Users can provide run time estimates while submitting tasks to the fog or IoT device. User run time estimates are recorded by all workloads. Given that predictions are

made, the next natural step is to determine the accuracy of these predictions. The accuracy of predictions is calculated by using (3) by the Performance Evaluator.

$$accuracy = \begin{cases} 1 & \text{if} \quad P_{RT} = R_T \\ P_{RT}/R_T & \text{if} \quad P_{RT} < R_T \\ R_T/P_{RT} & \text{if} \quad P_{RT} > R_T \end{cases} \tag{3}$$

Where, P_{RT} is predicted run time and R_T is actual run time.

The accuracy of predictions achieved by both the user and the system is presented in Table 1.

Table 1: Comparison of user's prediction accuracy and system's prediction accuracy

# No. of tasks	Average User % Accuracy	Average System % Accuracy	System 90% Confidence Interval	System 95% Confidence Interval	User 90% Confidence Interval	User 95% Confidence Interval
1000	5.9	73.0	81.4	83.6	2.0	2.0
2000	9.8	79.8	84.7	87.9	8	10
3000	15.3	86.9	88.8	90.7	14.7	16.5
4000	20.0	90.2	91.9	93.3	20.8	23.4
5000	27.1	95.0	96.2	96.9	35.0	39.0

From the results presented in Table 1, it can be seen that the smallest accuracy of 73% was achieved for the workloads of 1000 tasks. The largest accuracy of 95% was achieved for the workloads of 5000 tasks. The remaining workloads have also shown very high accuracy. The accuracy of predictions was also compared with user provided run time estimates while submitting tasks to the fog. User run time estimates are recorded, and as shown in the results in Table 1, the accuracy of user run time estimates lies between 5.9% and 27.1%. It can be concluded therefore that user accuracy is very poor when compared with the achieved accuracy from the predictor.

Another metric of interest, in the fog context, is the number of tasks falling within a certain confidence interval. This interval is defined by the Chebyshevs inequality theorem [22] and can be calculated when a sample's mean is being used as a predictor. On the fog where deadline based scheduling is expected, this confidence interval gives a better idea to the user about the expected task completion time. Chebyshevs' theorem states that the portion of data that lies within k standard deviations to the either side of the mean is at least $1 - \frac{1}{k^2}$ of any data set [22] [23] and it can be expressed as (3):

$$c = \left(1 - \frac{1}{k^2}\right) * 100 \tag{4}$$

Where c is referred to as the confidence interval and k is the number of standard deviations. The Performance Evaluator calculates the predicted tasks falling within the 90% confidence interval and 95% confidence interval. The results of both of these calculations are shown in Table 1. The predictor performed better again, since 81.4% to 96.2% tasks were completed within a 90% confidence interval and 83.6% to 96.9% tasks were completed within a 95% confidence interval. On the other hand, for user run time estimates, only 2.0% to 35.0% tasks were completed within a 90% confidence interval and 2.0% to 39.0% tasks were completed within a 95% confidence interval.

Large number of tasks were executed, first using the existing scheme then the prediction FResS model. The prediction model were generated using queries logs of large number of tasks from CloudSim to evaluate the performance of the proposed model, resource predictions from ANN were generated and used to execute the incoming tasks. The time taken by FResS and the current scheme to execute tasks was measured.

The stages of executing tasks on the fog are explained in the below section, which start with submitting the task and end with the retrieving results after completion. These stages comprise four major components and are given below:

1. **Stage in**: transferring the task(s) to executable resource.
2. **Waiting time**: waiting in the queue for its execution turn.
3. **Run time**: the task is assigned for execution, which marks the start of the run time.
4. **Stage out**: transferring the completed task(s) to the originating node.

Stage in also called stage one, starts with transferring the task from its originating node to the executing node. Once the destination node receives all the required data, then the task is forwarded to the local task manager, which in turns puts it in the waiting queue. The queue wait time represent the time spent in the queue for its execution turn, it depend on the load of the executing node. The run time in stage three marks the start of task execution. In stage four after execution is completed, the task is returned to the originating node which is called the stage out phase. The total turnaround time is given in (2).

$$T_{TRT} = T_{SIT} + T_{QWT} + T_{RT} + T_{SOT} \tag{2}$$

where:

T_{TRT} = *Total Run Time*

T_{SIT} = *Stage In Time*

T_{QWT} = *Queue Wait Time*

T_{RT} = *Run Time*

T_{SOT} = *Stage Out Time*

To evaluate FResS a number of experiments we designed. The purpose of these experiments was to evaluate overall time performance by varying the number of tasks. The tasks response time outlined in Figure 8 show that the FResS model outperform the existing scheme, and the results were promising where the tasks response time was decreased by 33%, as shown in Table 2 showing significant time savings.

Since the fog users have a limited network bandwidth to communicate among the system, such a high communication cost

is not acceptable. In such environment, nodes within a region is provided with broader bandwidth whereas nodes across region is provided with narrow bandwidth. Since many nodes from different regions try to connect to each other through a narrow bandwidth, causing congested network traffic which in turn will affect the task execution time which is an important factor, therefore, carrying out some experiments by varying and expanding the network bandwidths was important in order to simulate the real world environment, the bandwidth between sites was varied in a set of experiments to evaluate the performance of both models.

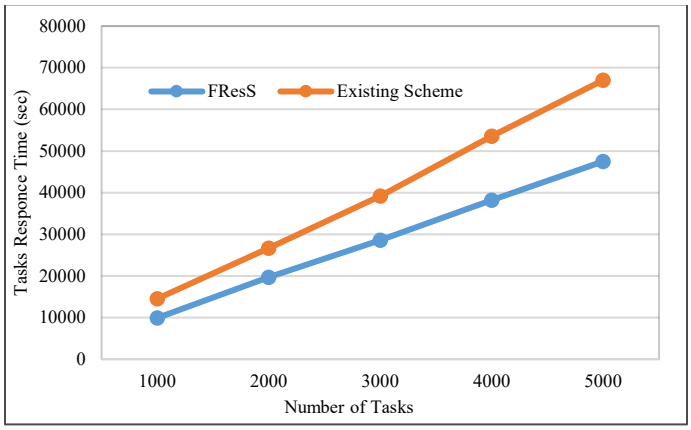


Figure 8: Comparative analysis of Response Time with no of tasks

Table 2. Simulation results: Task Turnaround Time using FResS and Existing scheme

# No. of tasks	Task Turnaround Time using existing scheme (Sec)	Task Turnaround Time using FResS (Sec)	Difference (Sec)
1000	14540	9920	4620
2000	26703	19710	6993
3000	39243	28608	10635
4000	53600	38214	15386
5000	67020	47512	19508
Total	201106	143964	57142
Average	28792	40221	33%

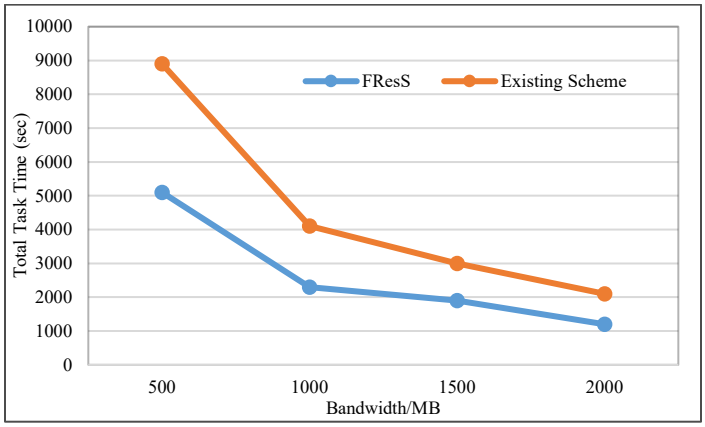


Figure 9: Total task time with varying bandwidth

Figure 9 shows the result of sets of experiments that were run on both narrow and broad bandwidth, as shown in the results, the FResS outperforms the existing scheme, even though setting broader bandwidth decreases the differences of task execution time, however, the difference is still significant in all scenarios. Consequently, it can be concluded that the cost and resource utilization are effectively considered in the proposed FResS model and outperform the existing scheme.

6. Conclusion and Future Work

to overcome the drawback of the cloud computing, the fog computing has emerged with the concept of sharing computational resources and information services by offloading the resources to the edge network to be closer to the devices that originate the requests rather than transporting the tasks to the distant cloud, thus reducing communication overhead, bandwidth consumption and latency. However, the complexity and dynamic nature of the fog systems demand a more coordination platform to handle such a large-scale numbers of resources, users and tasks requests. Therefore, the existing approaches, however, often exhibit a high cost in terms of response time and bandwidth consumption. These shortcomings were overcome by presenting a fully distributed FResS, which stores historical data related to a single user, which simplifies tasks management.

This paper proposed an extension to the resource selection service FResS based on neural network. The proposed prediction model minimize the total overhead of the task turnaround time of the incoming task by predicting and informing the system the required resources, and the database are updated constantly by storing the result of NN tool. The proposed prediction model is simple and shows comparatively lower overheads and provide high possible accuracy.

An evaluation strategy was devised to test five different algorithm used by the Resource Selector Module to provide a detailed performance analysis for each one. A History update algorithm was introduced to manage the history execution logs, as the size in IoT and fog environments are very limited. In addition, the experiments showed that the proposed prediction model distribute the load based on the user preferences. Another major strength of the proposed model is its effective resources utilization achieved by decreasing overall cost, response time, and bandwidth usage. In the future work, more metrics will be added to improve the load balancing, in addition to distributing the proposed model on the cloud layer. Moreover, more experiments will be carried out to define the size of the history records.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The writer would like to thank the American University of the Middle East for its support.

References

[1] N. Mostafa, "Cooperative Fog Communications using A Multi-Level Load Balancing," 2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC), Rome, Italy, 2019, pp. 45-51.doi: 10.1109/FMEC.2019.8795325.

[2] S. Zahara, I. Pratomo and D. S. Rahardjo, "Application and data level interoperability on virtual machine in cloud computing environment," 2015

1st International Conference on Wireless and Telematics (ICWT), Manado, 2015, pp.1-5. doi: 10.1109/ICWT.2015.7449238.

[3] V. B. Souza, X. Masip-Bruin, E. Marin-Tordera, W. Ramirez and S. Sanchez, "Towards Distributed Service Allocation in Fog-to-Cloud (F2C) Scenarios," 2016 IEEE Global Communications Conference (GLOBECOM), Washington, DC, 2016, pp. 1-6. doi: 10.1109/GLOCOM.2016.7842341.

[4] L. Ni, J. Zhang, C. Jiang, C. Yan and K. Yu, "Resource Allocation Strategy in Fog Computing Based on Priced Timed Petri Nets," in IEEE Internet of Things Journal, vol. 4, no. 5, pp. 1216-1228, Oct. 2017. doi: 10.1109/JIOT.2017.2709814.

[5] J. Yao and N. Ansari, "Task Allocation in Fog-Aided Mobile IoT by Lyapunov Online Reinforcement Learning," in IEEE Transactions on Green Communications and Networking. doi: 10.1109/TGCN.2019.2956626.

[6] S. Bian, X. Huang and Z. Shao, "Online Task Scheduling for Fog Computing with Multi-Resource Fairness," 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall), Honolulu, HI, USA, 2019, pp. 1-5. doi: 10.1109/VTCFall.2019.8891573.

[7] M. Mukherjee et al., "Task Data Offloading and Resource Allocation in Fog Computing With Multi-Task Delay Guarantee," in IEEE Access, vol. 7, pp. 152911-152918, 2019. doi: 10.1109/ACCESS.2019.2941741.

[8] F. Derakhshan, H. Roessler, P. Schefczik and S. Randriamasy, "On prediction of resource consumption of service requests in cloud environments," the 20th Conference on Innovations in Clouds, Internet and Networks (ICIN), Paris, pp. 169-176, 2017.

[9] Z. Zhang, B. Bockelman, D. W. Carder and T. Tannenbaum, "Lark: Bringing Network Awareness to High Throughput Computing," 2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, Shenzhen, pp. 382-391, 2015. doi: 10.1109/CCGrid.2015.116

[10] S. Yashpal, S. Alok, " Neural Network in Data Mining", Journal of Theoretical and Applied Information Technology, 37-42, 2009.

[11] C. Krieger, "Neural Networks in Data Mining", technician report, 1996.

[12] S. Duggal, R. Chhabra , "Learning Systems and Their Applications: Future of Strategic Expert System". Issues in Information Systems, Vol. III, 2002.

[13] G. jha, "Artificial Neural Networks," International journal of computer science and issues, Indian Research Institute, PUSA, New Delhi, 2005.

[14] S. Nissen, "Implementation of a fast artificial neural network library (FANN)," Technical report, Department of Computer Science University of Copenhagen (DIKU), 2003.

[15] D. Shanthi, G. Sahoo and N. Saravanan, "Designing an Artificial Neural Network Model for the Prediction of Thromboembolic Stroke," International Journals of Biometric and Bioinformatics (IJBB), Volume (3), 10-18, 2009.

[16] S. Chen, R. Lu, J. Zhang, "An Efficient Fog-Assisted Unstable Sensor Detection Scheme with Privacy Preserved", arXiv:1711.10190v1 [cs.CR] 28 Nov 2017.

[17] I. Rao and E. Huh, "A probabilistic and adaptive scheduling algorithm using system- generated predictions for inter-grid resource sharing," Journal of Supercomputer, 45, pp: 185-204, 2008.

[18] N. Moustafa, I. Al Ridhawi, and A. Hamza, "An Intelligent Dynamic Replica Selection Model within Grid Systems, " in Proc. 8th IEEE GCC conference on Towards Smart Sustainable Solutions, pp. 1-6, February 2015. doi: 10.1109/IEEEGCC.2015.7060061.

[19] "The Cloud Computing and Distributed Systems (CLOUDS) Laboratory, University of Melbourne," December. 2012, Available: http://www.cloudbus.org/cloudsim/.

[20] S. Mostafi, F. Khan, A. Chakrabarty, D. Y. Suh and M. J. Piran, "An Algorithm for Mapping a Traffic Domain Into a Complex Network: A Social Internet of Things Approach," in IEEE Access, vol. 7, pp. 40925-40940, 2019. doi: 10.1109/ACCESS.2019.2906647.

[21] B. Bhuyan, H. Sarma, N. Sarma, A. Kar and R. Mall, "Quality of Service (QoS) Provisions in Wireless Sensor Networks and Related Challenges", Wireless Sensor Network, Vol. 2 No. 11, pp. 861- 868, 2010. doi: 10.4236/wsn.2010.211104.

[22] W. J. Chu Using Chebyshev's Inequality to Determine Sample Size in Biometric Evaluation of Fingerprint Data, Forgotten Books, 2018.

[23] W. Smith, I. Foster, and V. Taylor, "Predicting application run times with historical information," Journal of Parallel Distributed Computing, pp. 1007–1016, 2004.

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 5, 409-418 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Big Data Security Layer Meta-Model Proposition

Allae Erraissi*, Abdessamad Belangour

Laboratory of Information Technology and Modeling, Hassan II University, Faculty of sciences Ben M'Sik, Casablanca, Morocc.

ARTICLE INFO

Article history:
Received: 11 July, 2019
Accepted: 09 October, 2019
Online: 28 October, 2019

Keywords:
Meta-model
Model Driven Engineering
Big Data
Security layer
Management layer

ABSTRACT

In Big Data, several solution providers offer distributions to handle this large amount of data. Given the variety of these solutions, we are working to provide universal meta-modeling for all layers of a Big Data system in order to address the issue of interoperability and portability between these solutions. As part of our continuous efforts to standardize concepts in Big Data world, we apply in this paper techniques related to Model-Driven Engineering "MDE" to propose a meta-model for the security layer in Big Data. This meta-model with the others that we have already proposed for the other layers of the Big Data system, will be used as a platform-independent according to Model-Driven Architecture pattern, which describes the structures of Big Data layers independently of any specific platform.

1. Introduction

The issues of the expression and application of security needs are present in all information systems. With the emergence of new types of environments, such as cloud computing and big data, this problem is becoming more complex. Hence, it is necessary to take into account their heterogeneity and the different levels of such an architecture to deal with this problem.

At the Big Data level, several distribution providers have proposed solutions to manage this huge amount of data, among these distributions, we found HortonWorks, Cloudera, MapR, InfoSphere BigInsghits IBM, Pivotal HD, etc. To put it another way, this massive amount of heterogeneous data led to the emergence of a large number of big data systems and technologies that share similar architectures but with different implementations. In essence, the common architecture of these data systems is composed of many components: Data sources, Ingestion, Hadoop Storage, Hadoop Platform management, Visualization, Monitoring, and Security Layers [1]. In our way for a unified abstract implementation, we proposed, in previous works, meta-models for data sources, ingestion [2], storage [3], management [4] and visualization layers [5]. We also relied on our previous comparatives studies to define key concepts of security layer in Big Data [6].

In this paper, we shall first present the security properties that express the security requirements of the software architecture. These properties are then combined to form a security policy: different models and languages of expression of security and insurance policies will be detailed. Then, we shall present the meta-model that we proposed for the security layer. This meta-model with the others already proposed for the other layers of the Big Data system will be used like a platform-independent according to Model-Driven Architecture pattern [7], which describes the structures of Big Data layers independently of any specific platform.

2. Related Work

Several research studies have been developed to standardize the security layer concepts at the Big Data systems level. Among the foremost relevant approaches in security modeling are UMLSec [8] and SecureUML [9]. Yet, both approaches use UML extensions to specify security requirements. UMLSec was created to verify the formal security requirements specifications within the system design. SecureUML was used to illustrate the MDS approach. SecureUML is specific to role-based access control infrastructures. It shows how an MDE policy could be applied to code generation for any aspect of security.

In the literature, several MDS approaches have been proposed. For example, on a middleware platform that integrates the implementation of the Corba component model with the OpenPMF security framework, Reznik et al. [10] present an MDS solution for

*Allae Erraissi, Laboratory of Information Technology and Modeling, Hassan II University, Faculty of sciences Ben M'Sik, Casablanca, Morocco, Email: erraissi.allae@gmail.com

developing secure applications. In the MDS approach, another UML profile was developed to model access control policies. Lang and Schreiner [11] show how OpenPMF architecture can be used to translate a high-level security-related regulatory requirement into enforceable authorization rules.

Our proposal is different from other existing MDS approaches, which use templates to represent design decisions and implementation details for target platforms. The application of Model Driven Engineering techniques that provide meta-modeling to the security layer in Big Data will advocate a systematic MDS application process. This application separates security requirements specifications and design decisions related to their implementation. Thus, the reuse of models is facilitated by this separation.

However, ModelSec proposes an intermediate model conforming to a target platform meta-model to reduce the semantic gap between the security requirements and the software generated. This intermediate step favors the reuse of the transformations. Unlike the other approaches that create UML profiles, a DSL is defined to express the security requirements. It has been implemented by applying meta-modeling techniques. Besides, ModelSec is a generic approach, whereas most MDS approaches are specific and they focus on access control policies.

Yet, other approaches manage security requirements. These approaches are not aligned with MDS, and they focus primarily on eliciting security requirements rather than generating software objects from them. Among these proposals are Secure Tropos [12] and [13], which introduce the concept of antigoal. Yu et al. [14] use ontologies and Haley et al. [15] present a framework consisting of a set of activities designed to meet security requirements.

3. Security Properties and Policies

In this section, we describe methods for formalizing the different security needs. We first present the usual security properties and those that can be derived from them. Then, we shall take existing models to apply these properties, security policies and languages to define them.

3.1. Security properties

Security properties are the basis for expressing security requirements. The set of security properties is commonly seen as a set derived from three main properties: confidentiality, integrity, and availability (CIA: Confidentiality, Integrity, Availability). The exact interpretation of what these three properties imply depends on the context of use. However, their definition and application is an essential part of the safety assessment criteria, at both European [16] and international [17]. Several definitions of these properties exist in the literature [16,17,18]. We present a synthesis here.

3.1.1. Confidentiality

Confidentiality is about preventing unauthorized disclosure of information. It aims to prohibit unauthorized access to information. Accordingly, the property of confidentiality implies that the information is accessible only by certain entities, that is to say, that certain entities must not be able to obtain the information. This property is often used in sensitive environments, such as defense. It concerns both direct access and information transfer.

3.1.2. Integrity

Integrity means the fact that information cannot be unauthorized modified or deleted, whether during processing, storage or transfer of information. This property does not only concern voluntary modifications and deletions, but also accidental acts. Just as in the case of confidentiality, the integrity property specifies all the entities authorized to modify or delete information. By default, other entities cannot alter the information.

3.1.3. Availability

The availability property expresses the ability to access information or a resource. It is related to the reliability of a system since an unavailable system is a failing system. The temporal notion of access is relative to the field of application: access to information or service on a critical system (for example, in the medical field) must be done more quickly than on a non-critical system (for example, a website).

3.1.4. Insurance

The property triple of confidentiality, integrity, and availability is sometimes extended with other properties [19], such as the insurance property. The purpose of this is to provide evidence that the other properties have been applied and that they have the desired effect. The assurance of a property is the verification that this property has been properly applied. The level of insurance required depends on the system in question and must, therefore, be adapted according to the criticality of the system. In sum, the purpose of insurance property is to verify that the other properties of the policy have been applied and that this application has had the desired effect.

3.1.5. Derived properties

Privacy, integrity, and availability properties are the basic concepts of security. They can be used to define derived properties which are special cases, subsets or combinations of these basic properties. In this section, we describe some of these derived properties.

- PROCESS CONFINEMENT

Process containment has been defined by Lampson [20]. The problem of containment concerns the prevention of the disclosure by a service or a process of information considered as confidential by the users of this service. According to Lampson, one of the features needed for a process is that it should not disclose information and it must not store information. Indeed, if a process stores information and a user can observe this process, then there is a risk that the user can access the information. If the process does not store the information, then it cannot be disclosed. This process containment property can thus be seen as the isolation of the process from the rest of the system.

- AUTHENTICATION

The authentication property is a property that allows or denies access to information or service to entities. Authentication is the process of establishing trust in the identity of an entity [21]. This authentication property is essential to apply the properties seen previously. Indeed, it makes it possible to establish the identity of

a user, a service or a system, which is necessary to give him the appropriate rights.

3.2. Security Policies

A security policy is a set of properties that express the security needs of a system or a set of systems. Most security properties are based on resource access control. As a result, the concepts introduced in access control models can be transposed into a language to formally express the properties of the previous section. In this section, we shall first describe the main models of historical access control. Then, we shall present security policy and insurance policy expression languages, which express and apply security properties.

3.2.1. Historical models

As defined above, some security properties can be applied by access control mechanisms. An access control system is usually modeled using the following three elements:

- A set of topics that are the active entities of the system (for example, processes);
- A set of objects that are the passive entities of the system, on which the subjects can perform actions (files, sockets, etc.);
- A set of permissions that represent the authorized actions between a subject and an object (reading, writing, etc.), or between two subjects (sending a signal).

- **Control of Discretionary Access**

Discretionary Access Control (DAC) is the default historical model present on most operating systems. In this model, the management of access rights to a resource is left to the discretion of the owner of that resource. For example, on UNIX, the owner of a file can set the read, write, and execute rights for itself, for the members of the group that owns the file, and for all other users in the system.

An access control model can be represented as a matrix, where a line represents a subject, a column represents an object or a subject, and each element of the matrix represents a set of permissions of the subject on the object (or on the second subject). This model was formalized by Lampson [22, 23] using capacity lists and Access Control Lists (ACLs). Therefore, it proposes to indicate, in an array A, the set D of the protection domains (representing program execution contexts, that is to say, the subjects) on the lines, and the set X of the objects on the columns. Hence, Lampson defines the lists of capacities (definition 1) which establish the permissions of a domain d on all the objects o of the system. This is the set of actions allowed for each domain d.

Definition 1: List of Capabilities

Given a d∈D domain, the list of capabilities for domain d is all couples:

$$\mathbf{(o, A[d, o]), \forall o \in X.} \quad (1)$$

Afterworlds, Lampson defines access control lists (definition 2) that specify the set of permissions granted on an object for each domain in the system.

Definition 2: Access Control List (ACL)

Given an o∈X object, the access control list (ACL) for object o is the set of pairs:

$$\mathbf{(d, A[d, o]), \forall d \in D.} \quad (2)$$

However, this model is complex to update. For example, when adding a new user to the system, a complete row must be added to the A matrix. The Lampson model has therefore evolved to the HRU model. In the HRU model [24], discretionary access control is modeled from a matrix P containing all the subjects' rights to the objects. In this template, subjects can edit the access control matrix to create or delete topics or objects, as well as edit existing permissions. The HRU model models protection using the following elements:

- An access control matrix P;
- A set S of subjects and a set O of objects;
- A set R of generic rights (reading, writing, execution, possession, etc.);
- A finite set C of commands representing all the operations provided by the operating system (creation of files, modification of rights, etc.);
- A set E of elementary operations: enter and delete (add and remove rights), create subject and create object, destroy subject and destroy object (destruction of subjects and objects).

A triplet (S, O, P) represents the configuration of the system protection. In order to study the problem of the safety of a protection system, the authors of the model HRU are interested in the transfer of privilege (right) occurring when a command inserts a right r in the matrix P. The safety problem can be defined in the following way: given an initial configuration of the security policy, a system is considered safe for a right r if none of the commands of this system causes the transfer of the right r. The authors then showed that if the commands contain only one elementary action, the safety problem is decidable but its verification algorithm is NP-complete. Nevertheless, in the general case (the commands contain several elementary actions), the problem is undecidable. In the case of an operating system, the commands are not mono-operational and the problem of the safety of the protection system is therefore undecidable. This shows that it is impossible to guarantee security properties with a discretionary access control model.

Other DAC models have been defined to extend the HRU model, including TAM (Typed Access Matrix) [25] and DTAM (Dynamic Typed Access Matrix) [26]. TAM extends the HRU model by incorporating a strong typing concept [27] which corresponds to the association of immutable security types to all the subjects and objects of the system. DTAM extends the TAM model by adding the ability to dynamically modify object types. The different models of discretionary access control are historical models. They are mainly used for system rights management.

- **Mandatory access control**

Discretionary access control leaves the management of resource permissions to their owners, that is, users of the system.

In practice, this system is limited since many attacks against systems aim to obtain privileged access (root, or super-user). Such an attack, so-called elevation of privilege, is intended to obtain rights greater than those possessed. Thus, it allows the user to gain full access to the system and its resources and to overcome discretionary control. Besides, studies [28,29] have shown that DAC models are vulnerable, particularly because of the need for users to correctly define the set of resource permissions. Any definition error can create a security vulnerability that can be exploited to gain privileges. Mandatory Access Control (MAC) is consequently intended to address this problem by imposing a security policy on system users. Anderson proposes the use of a reference monitor [30] to control the interactions between subjects and objects and to determine which ones are valid (that is, the ones allowed by the policy). This section shall present the main models of mandatory access control. These models explain the concepts of security properties (integrity, confidentiality) in order to apply them.

The Bell-LaPadula (BLP) model [31] is based on the confidentiality needs of the military and is accordingly intended to prevent disclosures. This model extends the HRU model by adding the notion of a label associated with each subject and object of the system. A label corresponds to a security level and is composed of two security identifiers: the first identifier, hierarchical, indicates the level of classification (for objects) or authorization (for subjects), e.g., secret or top secret. The second identifier, called category, specifies the organizations using the information, for example, military or private.

In addition to the classic rules which are defined by an access control matrix, two new rules are defined:

- **ss-property** (simple security property): for reading access to be authorized, the subject requesting it must have a level of authorization greater than or equal to that of the object;
- ***-property** (star property): Information can only be transferred from a lower classification object to a higher classification object.

These two rules make it possible to ensure the confidentiality of the information and its non-disclosure. However, the existence of hidden channels can cause information flows that cannot be controlled. For this reason, a more restrictive version of BLP has been proposed with the following rules:

- No Read Up: A subject requesting read access to an object must have a security level greater than or equal to the object;
- No Write Down: A subject requesting write-only access (adding data) to an object must have a security level that is less than or equal to the object.

Consequently, a subject requesting read and write access to an object must have the same level of security as the object. This model is sometimes called MLS (Multi-Level Security), and thus refers to the level system used to define security rules.

A dual model at BLP, Biba [32] has been defined to meet integrity needs. The Domain and Type Enforcement (DTE) model [33] is a high-level mandatory access control model. Unlike models such as BLP or Biba, it is not intended to apply a specific security property, but rather to define the allowed access between different entities in the system. Nonetheless, this model can serve as a basis for implementing security property models, such as BLP, Biba, or process containment. The DTE replaces notions of topics and objects with those of domains and types. Thus, each object has a type and each subject runs in a domain. Access rights on types and domains are then defined for each domain. Consequently, this model aims to restrict the resources accessible by a process (even for privileged processes by the principle of least privilege) and to control which processes have access to sensitive resources. The different models of mandatory access control can be applied at the same time at the system level, in distributed environments such as the case of military infrastructures, or even be integrated within a software.

3.2.2. Expression languages of security policies

The access control models that were presented above are the historical models that allowed the definition of security policy expression models and languages. A security policy consists of a set of rules that may have the purpose of applying properties, possibly following one of the described models or reusing certain concepts. In this section, we shall describe the main models and languages of a policy expression.

RBAC [Role Based Access Control] is a role-based access control model that simplifies the writing and management of a security policy. The administration of a mandatory access control policy involves the management of multiple access rules. Yet, this process is time-consuming and subject to errors. With RBAC, rules can be simplified by expressing them according to the role of the subject and not his or her identity.

Thus, within an organization, roles are defined from the functions of the different positions and permissions are assigned to these roles. Users have roles that allow them to obtain permission. Considering that the roles are not directly assigned to the users, their management is facilitated. For example, when adding a user, it is necessary to assign the corresponding roles.

Different versions of RBAC have been defined [34,35]:

- Core RBAC understands the basic concepts of RBAC and specifies that user-role and role-permission associations are of many-to-many types;
- Hierarchical RBAC adds inheritance support between roles: for example, a junior employee has a junior role with permissions and the senior role inherits permissions from the junior role;
- Constrained RBAC introduces constraints to apply the privilege separation principle. Since 2004, RBAC is a NIST standard [36].

OrBAC [37] (Organization Based Access Control) is an access control model based on the concept of organization. OrBAC abstracts the notions of subject, action, and object by those of role (as in the case of RBAC), activity and view. Subsequently, this makes it possible to group entities according to the security rules that concern them, and thus simplify the expression of the policy. In order to facilitate the expression of dynamic policies, OrBAC uses contexts. Three types of rules are possible: First, permissions which allow a role to perform an activity on a view, in a given

context. Second, bans when an activity is prohibited and finally, obligations when a role must perform an activity on a view in a given context. An OrBAC rule can be expressed as follows: A role can have permission, prohibition or obligation to perform an activity on a given view when the associated context is verified.

ABAC [38] (Attribute Based Access Control) is an access control model in which rights are given to users based on attributes. Attributes can be seen as features of system elements. An attribute consists of a type (for example, a role, a project, a sensitivity level, etc.) and a value that can be single-valued or multi-valued. For example, a role identifier or the value of the sensitivity level, etc.). These attributes can be compared to each other or to fixed values, which allows defining the rules of policy. For example, a rule evaluates one or more attributes of a subject to decide whether to allow or deny access to an object. Note that ABAC can be used with the DAC, MAC and RBAC models [39].

XACML [40] (eXtensible Access Control Markup Language) is a standard of the OASIS consortium [41]. It is both an access control policy definition language and a template for interpreting (allowing or disallowing) access requests based on this policy. XACML is based on the XML language. A set of XACML policies consists of policies, which are themselves composed of a set of elements: target, rules, a combination algorithm, and obligations. The target element indicates whether a policy should be applied to a given query. The rules specify the allowed accesses: they consist of a condition that is to say a Boolean expression. If this condition is verified, it causes an effect (authorization or prohibition) on the access. The rule combination algorithm defines how the policy should be interpreted when multiple rules may apply, for example, "permission overrides prohibition". Obligations are optional and allow you to perform an action when a rule is encountered, for instance, generate an alert. However, because of the flexibility and expressiveness of XACML, the definition of security policies can be complex [42]. XACML cannot easily allow the cloud user to express his own security needs.

Ponder [43] is a specification language for security and administration policies for distributed systems. It is a declarative and object-oriented language that supports different types of policies: authorization policies (specifying allowed or forbidden access for a subject), obligation (actions that a subject must perform in response to an event), restriction (actions that a subject should not perform), delegation (actions that a subject may delegate to another subject), constraint (to limit the application of other policies, for example in function of time). Meta-policies can also be defined to manage interactions between policies. Ponder2 [44] reuses Ponder's concepts by adapting them to autonomous systems. Hence, Ponder2 is based on a decentralized architecture, made up of self-managed components (SMCs) capable of interpreting and applying the policy on system objects. The objects considered by Ponder2 are Java objects that can communicate with SMCs. These objects must be adapted to be managed by Ponder2.

3.2.3. Expression language of insurance policies

Security policies may be supplemented by insurance policies. An insurance policy must express methods for assessing the level of protection of a system.

XCCDF [45] is an XML-based standard for specifying security checks and performance tests. Hence, it is, a language used for defining an insurance and verification policy. Besides, XCCDF can be used to automate vulnerability checks and responses when these vulnerabilities are detected. XCCDF is used by SCAP [46] (Security Content Automation Protocol), which is a set of specifications defined by NIST to standardize the format in which software vulnerabilities and security configurations are expressed. A distributed version of XCCDF, called DXCCDF [47], is used to express distributed vulnerabilities.

OVAL [48] is another standard aimed at unifying the expression of insurance policies. OVAL uses XML to evaluate the state of a system by performing a series of tests on the machine. The evaluation of the state of the system with OVAL is broken down into three stages: the representation of the system information, the description of the different states and the transmission of the results of the evaluation. The distributed version of OVAL is named DOVAL [49] and can handle distributed vulnerabilities.

A-PPL [50] is a policy language for expressing accountability obligations. A-PPL can be used for privacy management policies, access control, and usage control policies. Besides, A-PPL manages elements specific to the insurance: it makes it possible to define alerts in the event of a detected error and determine rules of localization of the data. It can also specify which are the characteristics wanted for the audit and the storing error messages.

4. Meta-model for Security layer:

Big Data solutions are more and more used. Indeed, many providers of Big Data solutions have already proposed distributions like HortonWorks, Cloudera, MapR, etc. Although each of these distributions has its vision for a Big Data system, they all share their necessary needs of the security layer. Accordingly, Security is a crucial element for both customers and service providers.

However, before presenting the Security layer, we deem it necessary to talk firstly about our latest contribution of meta-models. In fact, in our previous work, we have proposed meta-models for layers: Data Sources, Ingestion, Hadoop Storage [51], Hadoop Platform Management [52], and visualization of Big Data architecture. Accordingly, we realized that there is a direct relationship between the security layer and the other layers in Big Data. In order to show the link between the layers of the Big Data system, we shall present in figure 1 the following meta-package diagram. It clearly indicates the meta-packages: IngestionPkg, DataSourcesPkg, HadoopPlatformManagementPkg, VizualisationPkg, SecurityPkg, and MonitoringPkg and the dependency relationship that links them.

We present in the figure 2 the meta-model that we proposed for the Security layer in Big Data. This meta-model defines specific security concepts. To separate the access control mechanisms from the rest of the concepts, we have devised our meta-model to two separate parts: the meta-model of security concepts and the meta-model that defines access control mechanisms. The purpose of this separation is to offer the possibility of extending the meta-model of access control mechanisms in our future work. It should be mentioned that this extension of the meta-model will not affect the rest of the meta-model. The figure 2 shows the meta-model of the general concepts of the security layer.

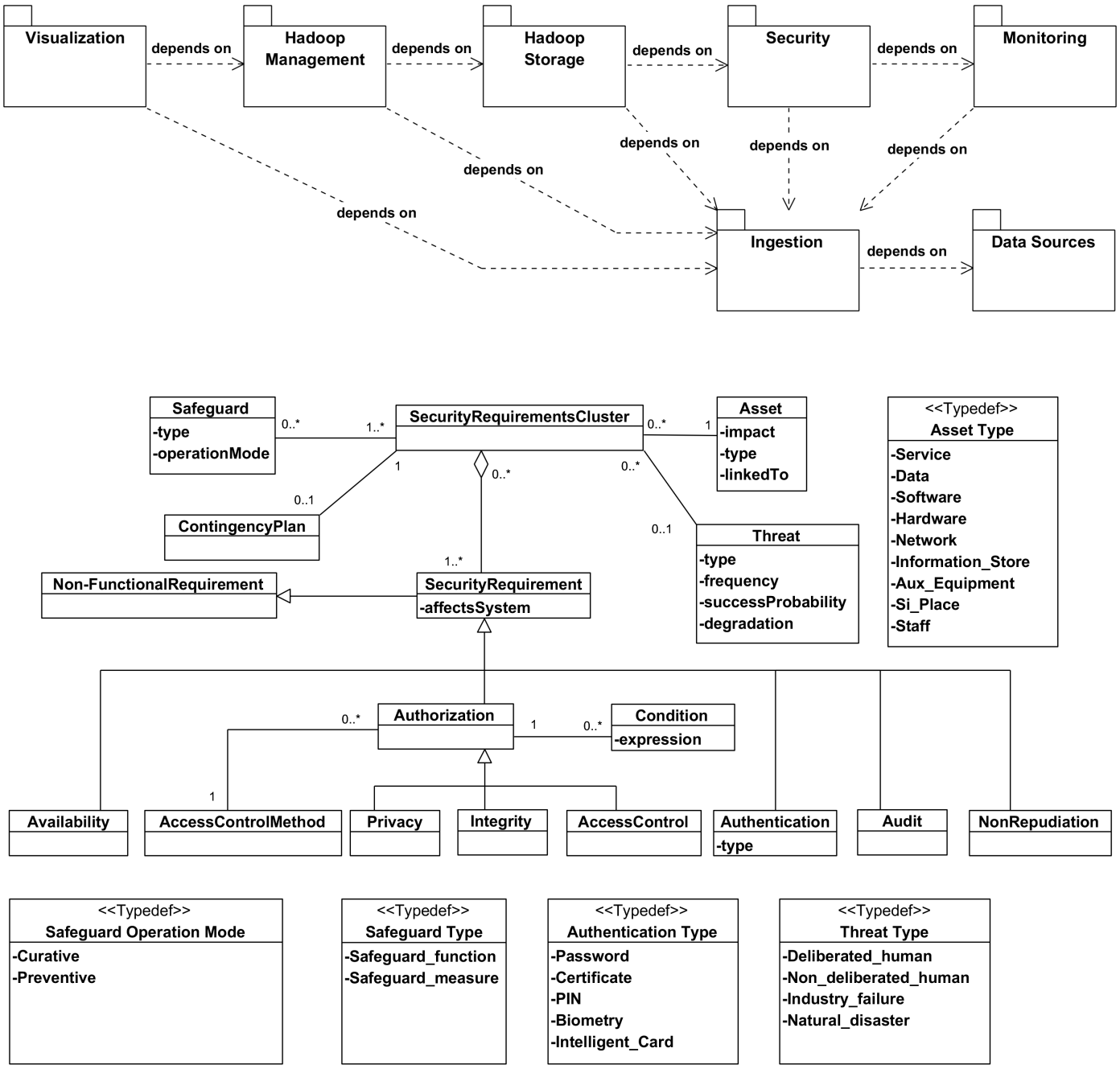


Figure. 2. Meta-model of general security concepts.

A threat can damage the assets. A threat has the following properties: type, frequency (modeled as an annual rate), probability of actual success, and degradation (that is, the level of damage to an asset if a threat reaches its goal).). Safeguards are a barrier to risk to reduce it. As indicated in the Safeguards type attribute, we can distinguish between Safeguard functions and Safeguard measures. Safeguards functions are actions that reduce risk while Safeguards' measures are physical or logical devices or

processes that reduce risk. Safeguards can act in two different modes, healing if they act on damaged assets or preventative if they act before a threat has appeared. A detailed emergency plan which consists of a set of safeguards is recommended to reduce a threat that may cause harm.

There is no standard classification for security requirements. According to [53], seven categories were taken into account (confidentiality, access control, integrity, authentication, non-repudiation, availability, and audit). Integrity is the guarantee that the information remains complete and correct. Access control is employed to force authorized users to access an asset. Confidentiality is to ensure that only authorized information can be read by those who are authorized. These three types of requirements are related to authentication and may have a condition that is defined as an expression. Authentication is a procedure by which a computer system certifies the identity of a person or a computer. The purpose of this procedure is to allow the person to access certain secure resources. It will compare the information of authorized users stored in a database (locally or on an authentication server) to those provided. Access will only be allowed if the information is the same. It is the administrator of the information system who grants the rights and sets the access. The user with an access account (ID + password) will only have access to the resources he is allowed to see. Availability maintains the proper functioning of the information system and ensures access to a service or resources. Finally, the audit assesses the IT risks of physical security, logical security, change management, emergency plan, etc. It also assesses a set of IT processes - which is usually the case - to respond to a specific customer request. All of these kinds of requirements can affect particular assets or the entire system.

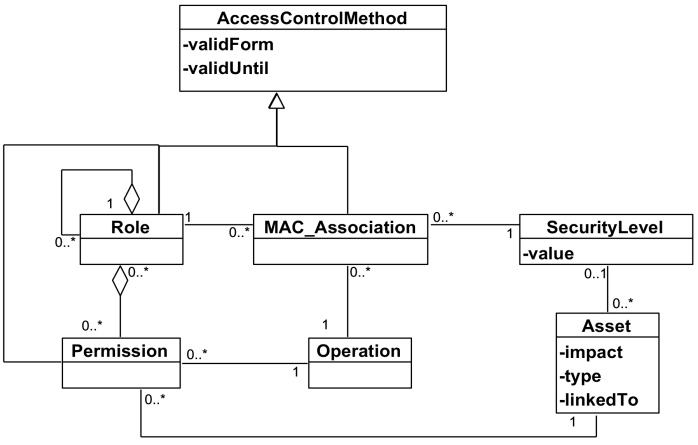


Figure. 3. Meta-model of access control mechanisms.

These two meta-models that we have proposed show the most important attributes, concepts, and data types that a Big Data system would need. Indeed, the meta-class of non-functional requirements is the extension point of the main meta-model and has the ability to add new non-functional requirements. The security requirement is a non-functional requirement specialization intended to be the root of meta-classes representing security concepts.

5. Transformations

The meta-models that we have defined for the different Big Data layers represent the platform-independent model (PIM) level, which means a model-independent of the platform. The goal of these meta-models is to standardize concepts at the Big Data level and to create an independent meta-modeling of platforms and solutions. Following that, we made some transformations using the Atlas Transformation Language (ATL). This transformation language will allow us to move from the PIM model to the Platform Specific Model (PSM). PSMs can use domain-specific languages or general languages like Java, C #, Python, and so on. The techniques used in the MDA approach are thus mainly modeling techniques and model transformation techniques. The figure below shows the transformations that we have made in this paper to make the transition from our meta-model of the security layer at the level of Big Data to PSM:

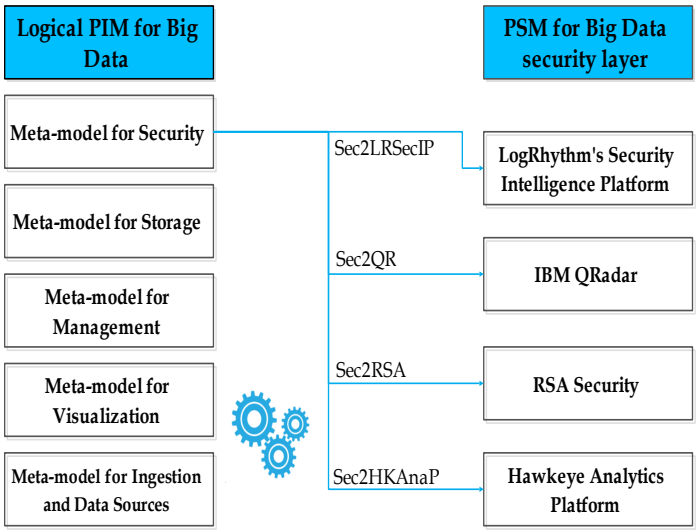


Figure. 4. Transformation of security meta-model to PSM.

6.1. Configuration

In this section, we shall discuss the techniques that we have used to implement the approach presented in Figure 4. Version 4.12 of the Eclipse IDE was used, with the addition of the Eclipse Modeling Framework (EMF) to draw the proposed meta-models. We also used version 4.1 of the ATL transformation language on

Eclipse 4.12, to define the transformation rules and to create our four transformations: Sec2LRSecIP, Sec2QR, Sec2RSA, and Sec2HKAnaP. The switch to the PSM did not take into consideration the version of the chosen solutions since the meta-model we proposed is a standard for the security layer. After applying the transformation rules, the result found will be used on all versions of LogRhythm's Security Intelligence Platform, IBM QRadar, RSA Security, and Hawkeye Analytics Platform.

6.2. Experiences

While performing our tests, we used two datasets to better measure the transformation execution time to the four chosen solutions: LogRhythm's Security Intelligence Platform, IBM QRadar, RSA Security, and Hawkeye Analytics Platform. The results are shown in Figures 5 and 6.

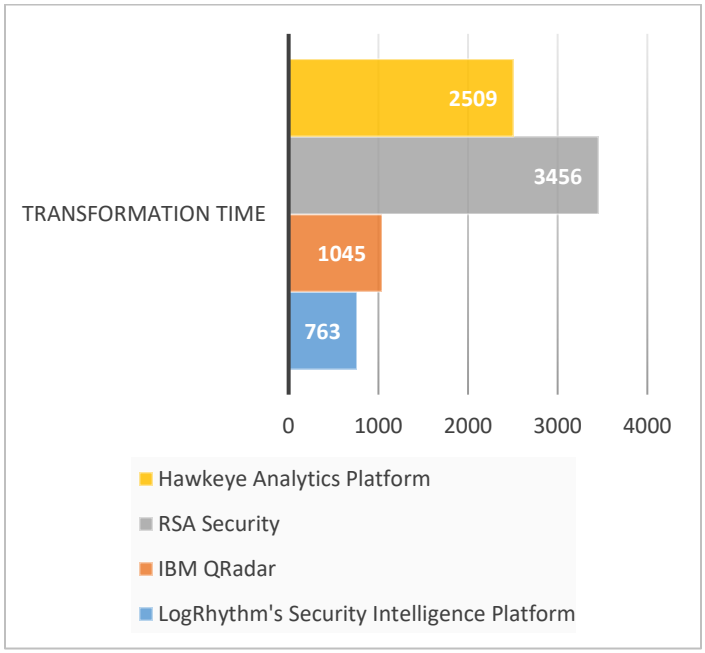


Figure. 5. Transformations time.

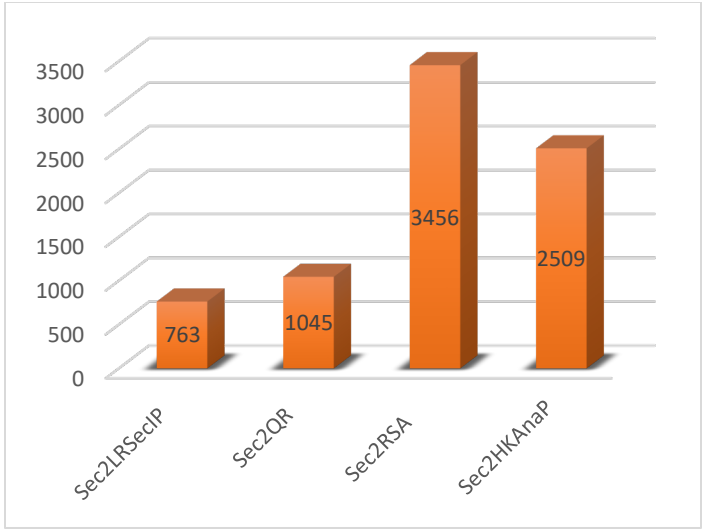


Figure. 6. Execution time for ATL transformations.

Table 1: Transformations runtime.

Security solution	LogRhythm's Security Intelligence Platform	IBM QRadar	RSA Security	Hawkeye Analytics Platform
Transformation time	763	1045	3456	2509

7. Discussion and Perspectives

The security properties that were presented in the first section are textual and abstract. They allow a user literally to express his security needs. Although these properties are high-level requirements, they cannot be directly applied to the system. This is why different models and security languages are designed to express these properties so that they can be applied to a system.

In the second section, we presented historical models for access control policies and security and insurance policy expression languages. Indeed, Historical access control models introduce concepts. These concepts can be generalized to security policies that are not dedicated to access control. For example, modeling in subject-permission-object triplet form is not necessarily specific to access control. Besides, access control has shown the value of a mandatory security policy compared to a discretionary policy. In fact, a discretionary policy gives the possibility to guarantee security properties [56,57]. Existing policy models and languages also make it possible to introduce essential notions into the expression languages of security and insurance policies. Thus, the notion of role introduced in RBAC is used in other models of security policies, such as OrBAC and XACML. It can simplify the expression and administration of policy. These languages allow the definition of security needs in the form of properties. However, these policies are often restricted to the expression of a property and support only a subset of properties. Moreover, existing languages are mostly not suitable for defining security and assurance properties for cloud and big data environments.

As we have mentioned before, Big Data has become essential in today's world. Indeed, our analysis of the subject leads us to find that there are several solutions to manage Big Data in the market (HortonWorks, Pivotal HD, IBM InfoSphere BigInsights, etc.). Each of these solutions manages Big Data in its own way without relying on standards like meta-models. This of course results in the diversity of solutions and the non-interoperability between the different solutions. During our research project, we have worked on the meta-modeling of the different layers of a Big Data system [58]. In this paper, we rely on comparative studies that we have already done to draw the key concepts of the security layer. The meta-model that we proposed for this layer and the other layers is platform-independent according to Model-Driven Architecture pattern. By using the ATL transformation language, for example, we can move to the other platform-specific models (PSM) [59,60]. The PSM we have chosen in this work are LogRhythm's Security Intelligence Platform, IBM QRadar, RSA Security, and Hawkeye Analytics Platform. After the creation of these meta-models, in the next step, we shall work on the creation of models respecting these meta-models. Then we shall define the transformation rules between these meta-models using the transformation language ATL (Atlas Transformation Language). These meta-models are platform-independent according to Model Driven Architecture

pattern, which describes the structures of Big Data layers independently from any specific platform.

8. Conclusion

Big data refers to the collection and aggregation of large amounts of data from different sources to extract information through statistical, descriptive and predictive analysis. Several new solutions exist in the IT market that can handle this huge amount of data. This big data trend makes security an essential but complex point to address. The definition of security needs can indeed be a difficult task, especially since the user of the service does not necessarily know these needs. However, we have seen that there are methods of risk analysis that allow a user to determine their security needs. Thus, we have presented in this paper the different basic security properties that may correspond to the needs of a user, as well as several models of policies to express these needs. After that, we have defined a standard meta-model for the security layer that represents the PIM level. Finally, we made transformations to move from our meta-model to four PSM.

References

[1] N. Sawant and H. (Software engineer) Shah, Big data application architecture Q & A a problem-solution approach. Apress, 2013.

[2] Erraissi, A., & Belangour, A. Data sources and ingestion big data layers: meta-modeling of key concepts and features. International Journal of Engineering & Technology, 7(4), 3607–3612, 2018. https://doi.org/10.14419/ijet.v7i4.21742

[3] Erraissi, Allae, and Abdessamad Belangour. « Hadoop Storage Big Data Layer: Meta-Modeling of Key Concepts and Features ». International Journal of Advanced Trends in Computer Science and Engineering 8, nº 3, 646-53, 2019. https://doi.org/10.30534/ijatcse/2019/49832019

[4] Erraissi, Allae, and Abdessamad Belangour. « Meta-Modeling of Big Data Management Layer ». International Journal of Emerging Trends in Engineering Research 7, no 7, 36-43, 2019. https://doi.org/10.30534/ijeter/2019/01772019.

[5] Erraissi, Allae, and Abdessamad Belangour. « Meta-Modeling of Big Data Visualization Layer Using On-Line Analytical Processing (OLAP) ». International Journal of Advanced Trends in Computer Science and Engineering 8, nº 4, 990-98, 2019. https://doi.org/10.30534/ijatcse/2019/02842019.

[6] Allae Erraissi, Abdessamad Belangour, and Abderrahim Tragha, "Digging into Hadoop-based Big Data Architectures," Int. J. Comput. Sci. Issues IJCSI, vol. 14, no. 6, pp. 52–59, Nov. 2017. https://doi.org/10.20943/01201706.5259

[7] Kleppe, A.G., Warmer, J., Warmer, J.B. and Bast, W., 2003. MDA explained: the model driven architecture: practice and promise. Addison-Wesley Professional, 2003.

[8] Jurjens, J. Secure Systems Development with UML. Springer Verlag, 2003.

[9] Basin, D., Doser, J., and Lodderstedt, T. Model driven security: From UML models to access control infrastructures. ACM Trans. Softw. Eng. Methodol., 15(1):39–91, 2006. https://doi.org/10.1145/1125808.1125810

[10] Reznik, J., Ritter, T., Schreiner, R., and Lang, U. Model driven development of security aspects. ENCTS, 163(2):65–79, 2007. https://doi.org/10.1016/j.entcs.2006.10.016

[11] Lang, U. and Schreiner, R. Managing business compliance using model-driven security management. In Securing Electronic Business Processes, pages 1–11, 2008. https://doi.org/10.1007/978-3-8348-9283-6_24

[12] Bresciani, P., Mouratidis, H., and Zanone, N. Modelling security and trust with secure tropos. Integrating Security and Software Engineering: Advances and Future Visions, pages 160–189, 2007. https://doi.org/10.4018/978-1-59904-147-6.ch008

[13] Van Lamsweerde, A. Elaborating security requirements by construction of intentional anti-models. In ICSE'04: 26th Int. Conf. on Software Engineering, pages 148–157. IEEE Computer Society, 2004. https://doi.org/10.1109/ICSE.2004.1317437

[14] Yu, E., Liu, L., and Mylopoulos, J. A social ontology for integrating security and software engineering. Integrating Security and Software Engineering: Advances and Future Visions, pages 70–109, 2007. https://doi.org/10.4018/9781605660608.ch048

[15] Haley, C., Laney, R., Moffett, J., and Nuseibeh, B. Security requirements engineering: A framework for representation and analysis. IEEE Trans. Software Eng., 34(1):133–153, 2008. https://doi.org/10.1109/TSE.2007.70754

[16] C.Jahl, ITSEC. Information Technology Security Evaluation Criteria (ITSEC) v1.2. Technical report, 1991. https://doi.org/10.1109/ICSE.1991.130656

[17] Tcsec, D. Trusted computer system evaluation criteria. DoD 5200.28-STD, 83, 1985.

[18] Bishop, M. What is computer security? Security & Privacy, IEEE, 1(1):67–69, 2003. https://doi.org/10.1109/MSECP.2003.1176998

[19] Stoneburner, G. Underlying technical models for information technology security: recommendation of the National Institute of Standards and Technology. US Department of Commerce, Computer Security Division, Information Technology, National Institute of Standards and Technology, 2001.

[20] Lampson, B. W. A note on the confinement problem. Communications of the ACM, 16(10):613–615, 1973. https://doi.org/10.1145/362375.362389

[21] Burr, W. E., Dodson, D. F. et Polk, W. T. Electronic authentication guideline: Recommendations of the national institute of standards and technology. NIST Special Publication, pages 800–63, 2013. https://doi.org/10.6028/NIST.SP.800-63-2

[22] Lampson, B. W. Dynamic protection structures. In Proceedings of the November 18-20, 1969, fall joint computer conference, pages 27–38. ACM, 1969. https://doi.org/10.1145/1478559.1478563

[23] Lampson, B. Protection. In Proc. 5th Princeton Conf. on Information Sciences and Systems, pages 18–24. Princeton, 1971. https://doi.org/10.1145/775265.775268

[24] Harrison, M. A., Ruzzo, W. L. et Ullman, J. D. Protection in operating systems. Communications of the ACM, 19(8):461–471, 1976. https://doi.org/10.1145/360303.360333

[25] Sandhu, R. S. The typed access matrix model. In Research in Security and Privacy. Proceedings. 1992 IEEE Computer Society Symposium on, pages 122–136. IEEE, 1992. https://doi.org/10.1109/RISP.1992.213266

[26] Soshi, M., Maekawa, M. et Okamoto, E. The dynamictyped access matrix model and decidability of the safety problem. IEICE transactions on fundamentals of electronics, communications and computer sciences, 87(1):190–203, 2004. https://doi.org/10.1007/10722599_7

[27] Sandhu, R. S. The schematic protection model: its definition and analysis for acyclic attenuating schemes. Journal of the ACM (JACM), 35(2):404–432, 1988. https://doi.org/10.1145/42282.42286

[28] Ferraiolo, D. et Kuhn, R. Role-based access control. In In 15th NIST-NCSC National Computer Security Conference, (1992). https://doi.org/10.1016/S0065-2458(08)60206-5

[29] Loscocco, P. A., Smalley, S. D., Muckelbauer, P. A., Taylor, R. C., Turner, S. J. et Farrell, J. F. The inevitability of failure: The flawed assumption of security in modern computing environments. In Proceedings of the 21st National Information Systems Security Conference, volume 10, pages 303–314, 1998.

[30] Anderson, J. P. Computer security threat monitoring and surveillance. Rapport technique, Technical report, James P. Anderson Company, Fort Washington, Pennsylvania, 1980.

[31] Bell, D. E. et LaPadula, L. J. Secure computer systems: Mathematical foundations. Rapport technique, DTIC Document, 1973.

[32] Biba, K. J. Integrity considerations for secure computer systems. Rapport technique, DTIC Document, 1977.

[33] Boebert, W. E. et Kain, R. Y. A practical alternative to hierarchical integrity policies. NIST SPECIAL PUBLICATION SP, pages A–10, 1989.

[34] Sandhu, R. S., Coyne, E. J., Feinstein, H. L. et Youman, C. E. Role-based access control models. Computer, (2):38–47, 1996. https://doi.org/10.1109/2.485845

[35] Sandhu, R., Ferraiolo, D. et Kuhn, R. The nist model for role-based access control: towards a unified standard. In ACM workshop on Role-based access control, volume 2000. https://doi.org/10.1145/344287.344301

[36] Standard, R. Incits 359-2004. ANSI INCITS, pages 359–2004, 2004.

[37] Kalam, A. A. E., Baida, R., Balbiani, P., Benferhat, S., Cuppens, F., Deswarte, Y., Miege, A., Saurel, C. et Trouessin, G. Organization based access control. In Policies for Distributed Systems and Networks, 2003. Proceedings. POLICY 2003. IEEE 4th International Workshop on, pages 120–131. IEEE, 2003. https://doi.org/10.1109/POLICY.2003.1206966

[38] Hu, V. C., Ferraiolo, D., Kuhn, R., Schnitzer, A., Sandlin, K., Miller, R. et Scarfone, K. Guide to attribute based access control (abac) definition and considerations. NIST Special Publication, 800:162, 2014. https://doi.org/10.6028/NIST.SP.800-162

[39] Jin, X., Krishnan, R. et Sandhu, R. S. A unified attributebased access control model covering dac, mac and rbac. DBSec, 12:41–55, 2012. https://doi.org/10.1007/978-3-642-31540-4_4

[40] Moses, T. et al. Extensible access control markup language (xacml) version 2.0. Oasis Standard, 200502, 2005.

[41] OASIS. OASIS. https://www.oasis-open.org/. 2015.

[42] Hu, V. C., Martin, E., Hwang, J. et Xie, T. Conformance checking of access control policies specified in xacml. In Computer Software and Applications Conference, 2007. COMPSAC 2007. 31st Annual International, volume 2, pages 275–280. IEEE, 2007. https://doi.org/10.1109/COMPSAC.2007.96
[43] Damianou, N., Dulay, N., Lupu, E. et Sloman, M. A language for specifying security and management policies for distributed systems. London: Department of Computing, Imperial College, Tech. Rep, 2000.
[44] Twidle, K., Dulay, N., Lupu, E. et Sloman, M. Ponder2: A policy system for autonomous pervasive environments. In Autonomic and Autonomous Systems, 2009. ICAS'09. Fifth International Conference on, pages 330–335. IEEE, 2009. https://doi.org/10.1109/ICAS.2009.42
[45] Waltermire, D., Schmidt, C., Scarfone, K. et Ziring, N. Specification for the extensible configuration checklist description format (xccdf) version 1.2 (draft). (2011b).
[46] Waltermire, D., Quinn, S., Scarfone, K. et Halbardier, A. The technical specification for the Security Content Automation Protocol (SCAP). NIST Special Publication, 800:126. (2011a).
[47] Barrère, M., Badonnel, R. et Festor, O. Collaborative remediation of configuration vulnerabilities in autonomic networks and systems. In Proceedings of the 8th International Conference on Network and Service Management, pages 357–363. International Federation for Information Processing. (2012a).
[48] OVAL Oval language. Http://oval.mitre.org/. 2014.
[49] Barrère, M., Badonnel, R. et Festor, O. Towards the assessment of distributed vulnerabilities in autonomic networks and systems. In Network Operations and Management Symposium (NOMS), 2012 IEEE, pages 335–342. IEEE. (2012b). https://doi.org/10.1109/NOMS.2012.6211916
[50] Azraoui, M., Elkhiyaoui, K., Önen, M., Bernsmed, K., De Oliveira, A. S. et Sendor, J. A-ppl : An accountability policy language. In Data Privacy Management, Autonomous Spontaneous Security, and Security Assurance, pages 319–326. Springer, 2015. https://doi.org/10.1007/978-3-319-17016-9
[51] A., Erraissi A., Belangour A. Capturing Hadoop Storage Big Data Layer Meta-Concepts. In: Ezziyyani M. (eds) Advanced Intelligent Systems for Sustainable Development (AI2SD'2018). AI2SD 2018. Advances in Intelligent Systems and Computing, vol 915. Springer, Cham, 2019. https://doi.org/10.1007/978-3-030-11928-7_37
[52] A. Erraissi and A. Belangour, "Meta-modeling of Zookeeper and MapReduce processing," 2018 International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), Kenitra, Morocco, pp. 1-5, 2018. https://doi.org/10.1109/ICECOCS.2018.8610630
[53] Rodriguez, A., Fern´andez-Medina, E., and Piattini, M. A bpmn extension for the modeling of security requirements in business processes. IEICE Transactions, 90-D(4):745–752, 2007. https://doi.org/10.1093/ietisy/e90-d.4.745
[54] Mellado, D., Fernandez-Medina, E., and Piattini, M. A common criteria based security requirements engineering process for the development of secure information systems. Comput. Stand. Interfaces, 29(2):244–253, 2007. https://doi.org/10.1016/j.csi.2006.04.002
[55] Samarati, P. and Capitani, S. D. Access control: Policies, models, and mechanisms. In FOSAD, pages 137–196, 2000. https://doi.org/10.1007/3-540-45608-2_3
[56] Ferraiolo, D. et Kuhn, R. Role-based access control. In In 15th NIST-NCSC National Computer Security Conference, 1992.
[57] Loscocco, P. A., Smalley, S. D., Muckelbauer, P. A., Taylor, R. C., Turner, S. J. et Farrell, J. F. (1998). The inevitability of failure: The flawed assumption of security in modern computing environments. In Proceedings of the 21st National Information Systems Security Conference, volume 10, pages 303–314.
[58] A. Erraissi, B. Mouad and A. Belangour, "A Big Data visualization layer meta-model proposition," 2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO), Manama, Bahrain, 2019, pp. 1-5. doi: 10.1109/ICMSAO.2019.8880276
[59] M. Banane, A. Erraissi and A. Belangour, "SPARQL2Hive: An approach to processing SPARQL queries on Hive based on meta-models," 2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO), Manama, Bahrain, 2019, pp. 1-5. doi: 10.1109/ICMSAO.2019.8880393
[60] Fatima Kalna, Allae Erraissi, Mouad Banane, Belangour "A Scalable Business Intelligence Decision-Making System in The Era of Big Data" International Journal of Innovative Technology and Exploring Engineering 2019. https://doi.org/10.35940/ijitee.L3251.1081219

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 411-418 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

A Comparative Analysis of ARIMA and Feed-Forward Neural Network Prognostic Model for Bull Services

Jude B. Rola[*,1], Cherry Lyn C. Sta. Romana[2], Larmie S. Feliscuzo[2], Ivy Fe M. Lopez[3], Cherry N. Rola[4]

[1]*Visayas State University, Department of Computer Science and Technology, Baybay City, 6521-a, Philippines*

[2]*Cebu Institute of Technology-University, Cebu City, 6000, Philippines*

[3]*Philippine Carabao Center-Visayas State University, 6521-a, Philippines*

[4]*Visayas State University, Department of Liberal Arts and Behavioral Sciences, Baybay City, 6521-a, Philippines*

ARTICLE INFO

Article history:
Received: 13 January, 2020
Accepted: 16 March, 2020
Online: 04 April, 2020

Keywords:
Time series forecasting
ARIMA
Artificial neural network

ABSTRACT

Bull service is the natural copulation by a purebred male carabao with a female counterpart. This is part of the bull loan agenda of the Philippine Carabao Center-Visayas State University (PCC-VSU), one of the 12 regional centers of PCC. For the past years, PCC-VSU used averaging of bull services count of previous years and sometimes did a little tweaking of the formulas in setting annual targets. Some targets fell short in the actual accomplishment that may have led to achieving lesser animal products. There is a need for resource optimization. Resource optimization through Auto-Regressive Integrated Moving Average (ARIMA) and Artificial Neural Network (ANN) forecasting can be used in the setting of goals of future number of copulations. This study deals with the comparative assessment of ARIMA and ANN in 4-step and 12-step ahead time series forecasting of bull services. This endeavor dealt with quarterly data from 2002 to 2017. In ARIMA, the for 4-step ahead (1-year) forecasting had the data from 2002 to 2016 for training while data from 2017 was employed for testing. ARIMA 12-step ahead (3-year) forecasting had data from 2002 to 2014 for training and 2015 to 2017 for testing. The ANN 4-step ahead prediction made use of data from 2002 to 2015 for training, 2016 for validation and 2017 for testing. ANN 12-step forecasting had data from 2002 to 2011 for training, 2012 to 2014 for validation and 2015 to 2017 for testing. R studio, a tool for statistical computing was employed in the analysis. Based on the performance measures Root Mean Square Error, Mean Absolute Percent Error and Mean Absolute Error the ANN outshines the ARIMA in both 4-step and 12-step ahead forecasting. The ANN forecasting can surely help in optimizing the center's resources. The forecasting of bull services using multivariate series is recommended.

1. Introduction

Bull service is the natural breeding by a purebred male carabao/water buffalo with a female mate. Dealing with the comparison of forecasting methods that predict bull services, this paper is an extension of the work originally presented in the 2019 International Conference on Multimedia Analysis and Pattern Recognition (MAPR) [1]. Bull service is among the services offered by the Philippine Carabao Center-Visayas State University (PCC-VSU) under the bull loan program which promotes water buffalo or carabao propagation and genomic enhancement.

PCC-VSU is among the 12 regional centers of PCC spread across the Philippine archipelago. The center is a government agency founded in 1992 through the Republic Act 7307. Other services it offers are artificial insemination (AI), buffalo semen distribution, and farmers and technicians' training [2]. PCC-VSU is mandated to deliver these services in region 8 comprising 6

[*] Jude B. Rola, Philippines, 63-53-563-7068/ jude.rola@vsu.edu.ph

provinces namely: Biliran, Leyte, Southern Leyte, Northern Samar, Samar, and Eastern Samar.

In the past, PCC-VSU used mean computation of bull services count of recent previous years and sometimes did little adjustment of formulas in setting annual targets. Some targets would fall short in the actual accomplishment and some would go over. Achieving more than the actual is an advantage but achieving low because of the low target does not optimize the animal and human resources. Forecasting bull services can help the center in the target setting and in optimizing its resources. It may help in aiming high to accomplish more. More bull services conducted means more animal products are produced – higher income especially for small-scale animal raisers. The focal purpose of this inquiry is to explore and compare the 4-step (1 year) and 12-step ahead (3 years) time series forecasting performance of Auto-Regressive Integrated Moving Average (ARIMA) and multi-layer artificial neural network (ANN) on bull services to improve the target setting competence of the center. In [1], only a single-layer ANN was used while in this research endeavor a multi-layer ANN was employed. This analysis utilized quarterly bull services data from 2002 to 2017. R studio was utilized in the analysis. The performance of ARIMA and ANN models were tested using the following measures: Root Mean Square Error (RMSE), Mean Absolute Percent Error (MAPE), and Mean Absolute Error (MAE).

In 2015, [3] combined ARIMA with Deep Belief Net (DBN) in predicting the residual of prediction results. Experimental results showed an effective forecast. Relatively, [4] produced a model to forecast the crude matters required for plastic products based in revenue data. The short-term prediction had relatively fine accuracy.

On the other hand, [5] established a neural network model with 4 input stratum neurons, 8 hidden stratum neurons and 1 output stratum neuron that effectively forecast information security incidents. [6] established an ANN model for forecasting and optimization of solar energy potential. Another model was produced in the study of [7] for wind speed forecasting. A recurrent neural network model was designed and implemented also by [8] for repairable system failure analysis. Another by [9] produced an ANN model that successfully predicted the rainfall one-step ahead. None of these tackled bull services prognostic modeling using a multi-layer ANN. A thorough web search was performed and none has been found as well hence, the purpose of the study.

2. Method

2.1. Box-Jenkins Method or Autoregressive Integrated Moving Average (ARIMA)

Autoregressive-Moving Average (ARMA) model incorporates a broad collection of non-stationary time series. It is then simplified producing an ARIMA. With the past data, the differencing technique is utilized to enhance stationarity into the modelling. The modelling decomposes the data into an autoregressive (AR) process, integrated (I) process, and a moving average (MA) of prognostic oversight.

The purpose of the AR procedure is to take a recollection of historical events, the I procedure is responsible for easing data stationarity, hence helps in easier prediction. Forecasting is more accurate with longer past data as it learns eventually. To achieve a transformation of a non-stationary procedure into a stationary one, a random walk is employed. A procedure $\{y_t\}$ is articulated to observe an Integrated ARMA model, conveyed by ARIMA (p, d, q), if $dy_t = (1–B)d\ \varepsilon_t$ is ARMA (p, q). The prognostic model can be written as

$$\varphi(B)(1-B)^d Y_t = \theta(B)\varepsilon_t \quad (1)$$

where

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_p B^q$$

$\varepsilon_t \sim$ WN (0, σ2), WN signifying White Noise. B indicates the backshift operator ($By_t = y_t –1$). The *d* is a non-negative integer that is known as the integration parameter. The ARIMA (p, d, q) model becomes an ARMA (p, q) model when d=0.

The 3 stages of ARIMA are identification, estimation and testing, and forecasting application. The constraints of the provisional chosen model at the identification part are approximated at the estimation stage while the model's adequacy is evaluated at the estimation and testing stage. When the model is unsatisfactory, the process is repeated until a sufficient model is selected [10]. Figure 1 depicts the ARIMA or Box-Jenkins modeling phases created by [11].

The Box-Jenkins Method or ARIMA is employed to forecast the trend of bull services for the next 4-quarters and 12-quarters. Its forecasts are then compared to the ANN's. The method that gives better prediction may help in determining the possible number of bulls and manpower that possibly maybe involve in carrying out bull services.

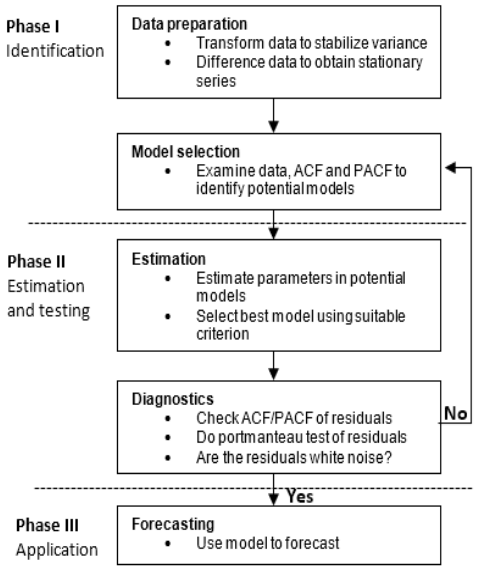


Figure 1: Box-Jenkins or ARIMA modeling approach.

2.2. Artificial Neural Network (ANN)

As great function estimators, the ANNs, commonly used in forecasting, pattern recognition, and control engineering replicate the human intellect. The Multi-Layer Perceptron (MLP) is an ANN of a feed-forward construct that is a compound layer network usually with 3 tiers namely: the input, the hidden, and the output layers. The MLP diagram, adapted from [3] is shown in Figure 2.

Each layer of the network has neurons or units. The neurons at the input layer obtain inputs from external entity and reproduce these inputs to the neurons of the hidden stratum. As the input signal components with the weights go beyond the biases, the neurons on the hidden layer can fire. A similar thing could happen among hidden and output layers. The neurons at the hidden and output layer generate function utilized logistic sigmoid depicted in (2) [3].

$$f(x) = \frac{1}{1 + exp\,(^{-x}/_{\varepsilon})} \qquad \varepsilon : \text{Gradient} \qquad (2)$$

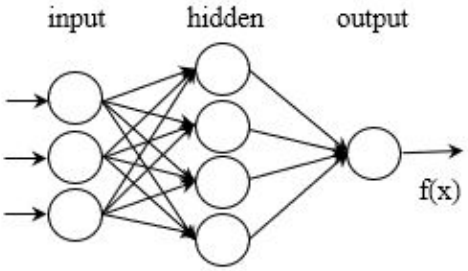


Figure 2: The structure of a multilayer perceptron.

2.3. Performance Evolution and Error Estimations

Some of the evaluation functions utilized in determining the success of the ARIMA method are Akaike information criterion (AIC), Bayesian information criterion (BIC), adjusted coefficient of determination (˙R), mean absolute percent error (MAPE), root mean square error (RMSE), and mean absolute error (MAE) [11, 12]. AIC and BIC select models with more accuracy and of low error in out-of-sample forecasting. The errors of these techniques are computed by the coefficient of determination and MAPE [12, 13, 14]. The MAPE is calculated by getting the difference between predicted and actual value. The outcome of subtraction is divided by the actual value, multiplied by 100, and then divided by the number of samples [15]. RMSE presents the concentration of data near the line of best fit. Aside from forecasting, it is widely utilized in regression and climatology analyses [16]. It is employed to measure the projection error values of potential models for a specific dataset [17]. MAE, on the other hand, entails adding the magnitudes of the errors and dividing it by n [18].

3. Experimental Results

3.1. Data and Modeling Tool

Figure 3 shows the time series chart of bull services covering 2002 to 2017. In ARIMA modeling, the for 4-step ahead (1-year) forecasting had the data from 2002 to 2016 (60 samples) for training while data from 2017 was employed for testing. ARIMA 12-step ahead (3-year) forecasting had data from 2002 to 2014 (52 samples) for training and 2015 to 2017 for testing.

On the other hand, ANN 4-step ahead prediction made use of data from 2002 to 2015 (56 samples) for training, 2016 for validation and 2017 for testing. ANN 12-step forecasting had data from 2002 to 2011 (40 samples) for training, 2012 to 2014 for validation and 2015 to 2017 for testing. R studio was the statistical computing tool utilized for both ARIMA and ANN modeling.

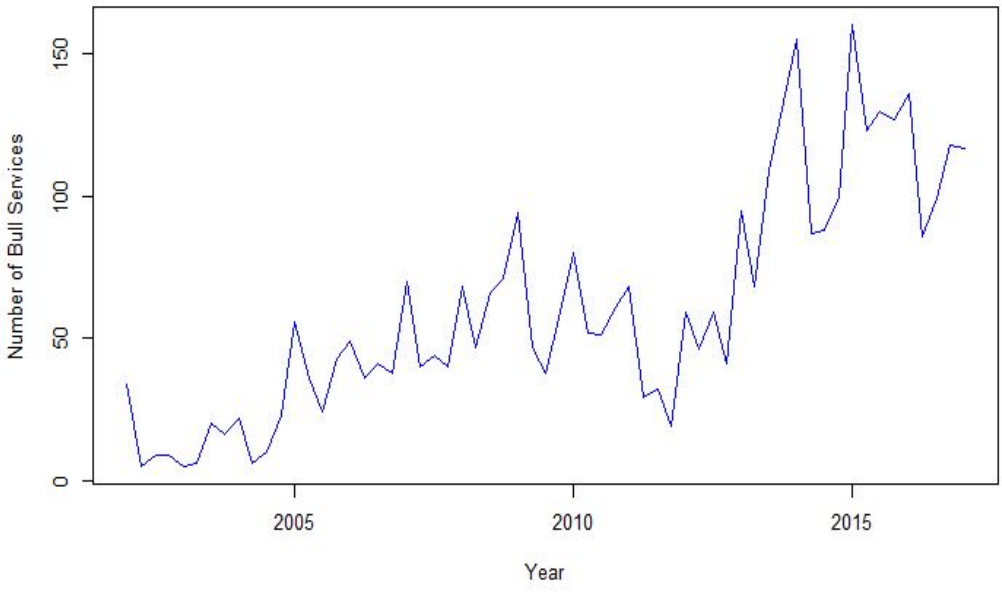


Figure 3: Time series plot of bull services (2002-2017).

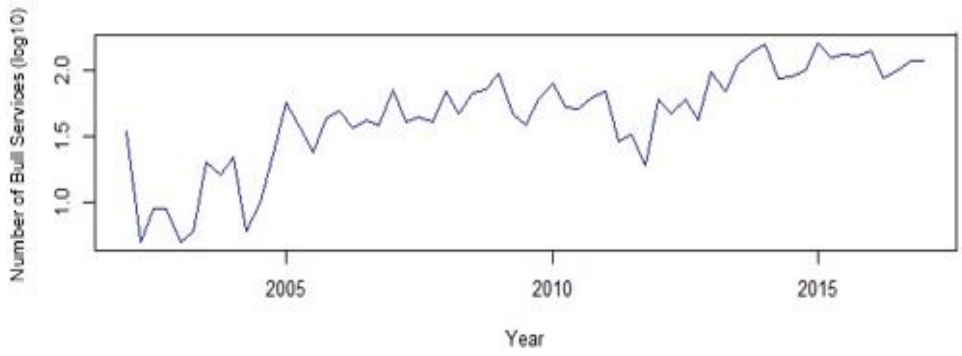


Figure 4: Logarithmic- transformation of bull services data series.

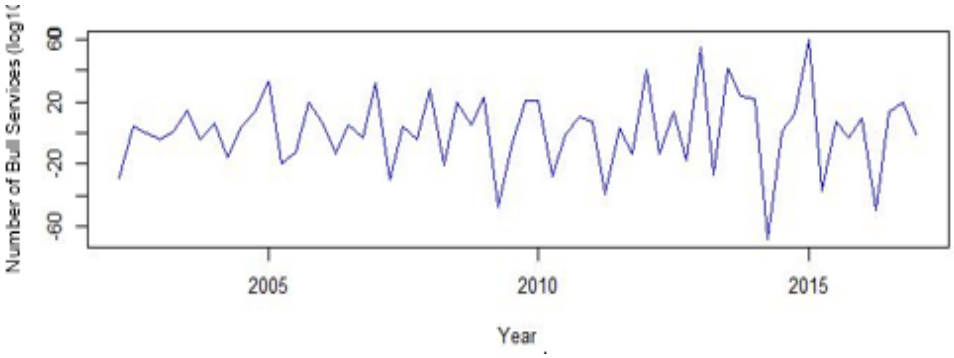


Figure 5: Differenced Bull Services Data

3.2. Preprocessing Data

The time series plot of bull services from 2002 to 2017 is shown in Figure 3. There is a growing activity inclination from the beginning to the initial quarter of 2009. During the sophomore quarter of the same year to the last quarter of 2012, a fall in the number of services can be observed. It starts to rise again in the initial quarter of the following data point. The primary quarter of 2014 is considered as its peak; however, 2014 falls again. Furthermore, the initial periods of 2005 to 2012 depicts the greatest number of bull services.

Generally, just like the other regional centers, along the years PCC-VSU may have met some challenges in the implementation of the bull services that were evident in the years 2010 to 2015 due to few significant fluctuations of data. The factors that may have affected it were the social and economic aspect of the bull loan beneficiary, willingness in adapting the process of acquiring bulls, PCC and LGU staff monitoring difficulties, political intervention in choosing recipients, and bull handlers' expertise deficiency. These factors were manifested in the study of [19].

There was a logarithmic conversion or transformation of data as shown in Figure 4 since there was no stationarity in the methods which predict series (Y_t) of raw data. The process makes data stationary on variance. The log transformation equation is depicted in (3). The transformed data were used in both ARIMA and neural network forecasting.

$$Y_t^{new} = \log_{10} Y_t \quad (3)$$

The data were then subjected to one-time differencing to transform its mean to stationary (on logged) (Figure 5).

3.3. ARIMA Model

For both the 4-step ahead (1-year) and 12-step ahead (3 years) bull services forecasting, the best model is ARIMA (1,1,1) (2,0,0) [4] based on AIC, BIC, and AIC_c. Table 1 and Table 2 presents the AIC_c values.

3.4. Feed-forward Neural Network Model

The suitable ANNs for both 4-step ahead and 12-step ahead bull services forecasts were chosen from various examined networks in the training stage; but, only sample results of different input parameters are presented on this paper (Table 3 and Table 4) employing the following number of epochs: 100, 500 and 1000. The learning rate used was 0.001. For shorter-term forecasting, the finest outcome was produced by the model which comprises 20 input neurons (window width), 12 hidden neurons and 4 output neurons trained for 1000 epochs (Table 3).

For longer-term forecasting, the best outcome was from the network which consists of 4 input neurons (window width), 4 hidden neurons and 12 output neurons trained for 100 epochs (Table 4). There was no overfitting that occurred in ANN modeling as the same network structures in both 4-step and 12-step ahead forecasting gave the most accurate prediction in the validation and testing stage.

Table 1: Statistical results of different ARIMA parameters for PCC-VSU bull services (4-step ahead forecast).

ARIMA	AIC_c
(0, 1, 0)	-5.43
(0, 1, 1)	-14.21
(0, 1, 1) (1, 0, 0) [4]	-15.76
(1, 1, 1) (1, 0, 0) [4]	-18.20
(1, 1, 0) (1, 0, 0) [4]	-13.43
(1, 1, 2) (1, 0, 0) [4]	-16.08
(0, 1, 0) (1, 0, 0) [4]	-14.10
(2, 1, 2) (1, 0, 0) [4]	-15.14
(1, 1, 1)	-12.54
(1, 1, 1) (2, 0, 0) [4]	**-19.87**
(0, 1, 1) (2, 0, 0) [4]	-17.25
(2, 1, 1) (2, 0, 0) [4]	-17.43
(1, 1, 0) (2, 0, 0) [4]	-15.98
(1, 1, 2) (2, 0, 0) [4]	-17.43
(0, 1, 0) (2, 0, 0) [4]	-16.60
(2, 1, 2) (2, 0, 0) [4]	-14.89

Table 2: Statistical results of different ARIMA parameters for PCC-VSU bull services (12-step ahead forecast).

ARIMA	AIC_c
(0, 1, 0)	4.60
(0, 1, 1)	-1.98
(0, 1, 1) (1, 0, 0) [4]	-2.35
(1, 1, 1) (1, 0, 0) [4]	-3.59
(1, 1, 0) (1, 0, 0) [4]	-0.39
(1, 1, 2) (1, 0, 0) [4]	-1.24
(0, 1, 0) (1, 0, 0) [4]	-1.15
(1, 1, 1)	-0.17
(1, 1, 1) (2, 0, 0) [4]	**-3.70**
(0, 1, 1) (2, 0, 0) [4]	-2.33
(2, 1, 1) (2, 0, 0) [4]	-1.13
(1, 1, 0) (2, 0, 0) [4]	-1.12
(1, 1, 2) (2, 0, 0) [4]	-1.13
(0, 1, 0) (2, 0, 0) [4]	-1.91

Table 3: Statistical performance of ANN model for bull services (4-step ahead forecast; learning rate =0.001).

Network Structure	**RMSE**	**MAPE (%)**	**MAE**	**RMSE**	**MAPE (%)**	**MAE**	**RMSE**	**MAPE (%)**	**MAE**
	100 epochs			**500 epochs**			**1000 epochs**		
20-3-4	0.14	6.91	0.13	0.04	1.83	0.04	0.07	2.66	0.05
20-4-4	0.14	6.87	0.13	0.15	6.22	0.12	0.15	6.00	0.12
20-5-4	0.17	8.21	0.16	0.15	7.49	0.15	0.14	5.87	0.12
20-6-4	0.13	6.33	0.12	0.16	8.00	0.16	0.15	7.28	0.14
20-7-4	0.14	6.91	0.13	0.15	6.67	0.13	0.16	7.28	0.14
20-8-4	0.15	6.76	0.13	0.14	7.16	0.14	0.14	6.10	0.12
20-9-4	0.15	7.20	0.14	0.13	6.07	0.12	0.15	7.32	0.14
20-10-4	0.16	7.51	0.15	0.18	8.45	0.17	0.14	6.70	0.13
20-11-4	0.15	6.54	0.13	0.17	7.78	0.15	0.16	6.93	0.14
20-12-4	0.16	7.63	0.15	0.13	6.48	0.13	**0.02**	**0.55**	**0.01**

Table 4: Statistical performance of ANN model for bull services (12-step ahead forecast; learning rate =0.001).

Network Structure	**RMSE**	**MAPE**	**MAE**	**RMSE**	**MAPE**	**MAE**	**RMSE**	**MAPE**	**MAE**
	100 epochs			**500 epochs**			**1000 epochs**		
4-3-12	0.11	4.01	0.08	0.08	3.30	0.07	0.09	3.86	0.08
4-4-12	**0.07**	**2.77**	**0.06**	0.10	3.88	0.08	0.09	3.70	0.08
4-5-12	0.14	4.44	0.09	0.15	6.14	0.13	0.19	7.46	0.15
4-6-12	0.13	5.29	0.11	0.15	5.62	0.12	0.12	4.70	0.10
4-7-12	0.12	4.78	0.10	0.08	3.05	006	0.13	5.64	0.12
4-8-12	0.13	4.77	0.10	0.20	7.84	0.16	0.18	6.88	0.14
4-9-12	0.18	7.65	0.16	0.09	2.99	0.06	0.22	8.57	0.18
4-10-12	0.16	6.09	0.12	0.19	6.81	0.14	0.25	9.79	0.20
4-11-12	0.18	6.64	0.14	0.50	20.67	0.43	0.24	9.20	0.19
4-12-12	0.34	6.39	0.25	0.25	10.11	0.21	0.21	7.35	0.15

3.5. Comparison of ARIMA and ANN Model

The comparative analysis utilized the testing datasets. In terms of performance error measures RMSE, MAPE, and MAE, the ANN model outperforms the ARIMA in both 4-quarter or 1-year bull natural mating forecasting (Table 5) although ARIMA is known as good in short-term forecasting in many cases [20]. The ARIMA model RMSE, MAPE, and MAE errors are 0.13, 6.45% and 0.12, respectively while ANN scored 0.02, 0.55% and 0.01. This clearly shows ANN to be better because its error values are smaller than the ARIMA. Figure 6 depicts a graph of ARIMA 4-step ahead forecasts while Figure 7 shows the ANN forecast (y-axis is on logged). The blue line exhibits the forecasted data and the red one represents the actual. Subscribing to the values above, forecasted data by ANN is closer to the target data than the ARIMA's (Figure 6 and 7).

For 12-step ahead or 3-year forecasting, ANN is still better than its counterpart. Table 6 shows the empirical outcome of ARIMA's RMSE, MAPE, and MAE scores of 0.08, 3.31% and 0.07 while ANN scores 0.07, 2.77% and 0.06. Figure 8 portrays a graph of the ARIMA 12-step ahead forecast whereas Figure 9 (y-axis is on logged) exhibits the ANN's. Still subscribed to the values above, forecasted data by ANN is closer to the target data than the ARIMA's (Figure 8 and Figure 9) since ANN got smaller error values than ARIMA.

RMSE is utilized since this study analyzes different models of the same dataset [17]. Conferring to [21], both models generated good models since the models scored lower than 10% of MAPE. MAE is employed as a common performance measure for time series forecasting [17]. Generally, the results subscribe to the study of [22] and [23] that ANN is better than ARIMA in forecasting.

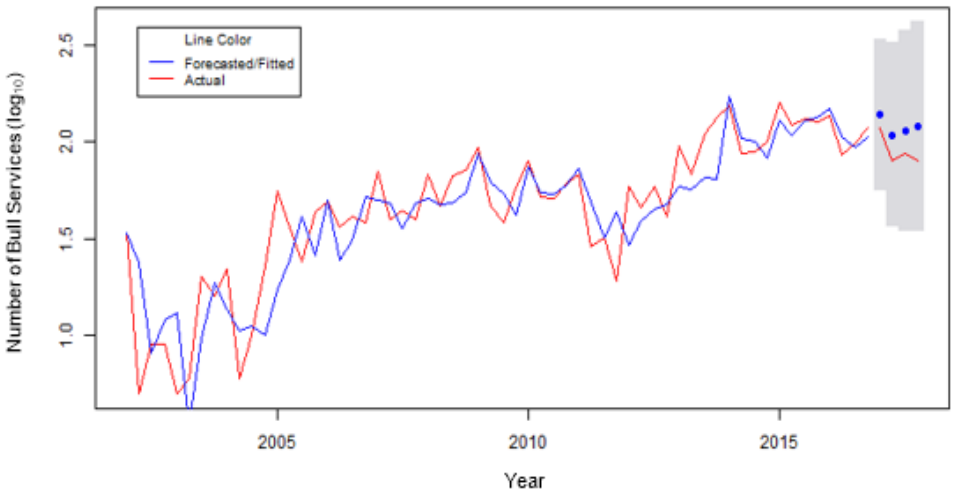


Figure 6: ARIMA 4-step ahead forecast.

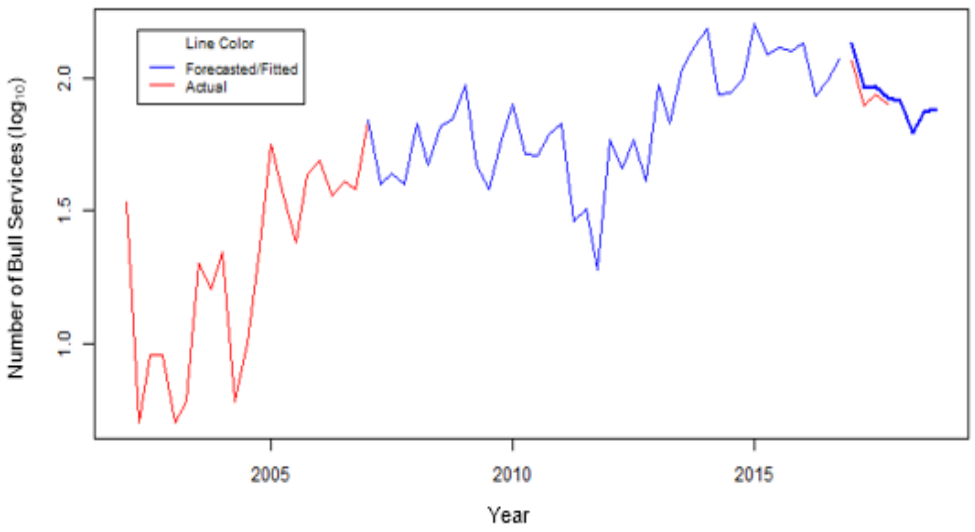


Figure 7: ANN 4-step ahead forecast.

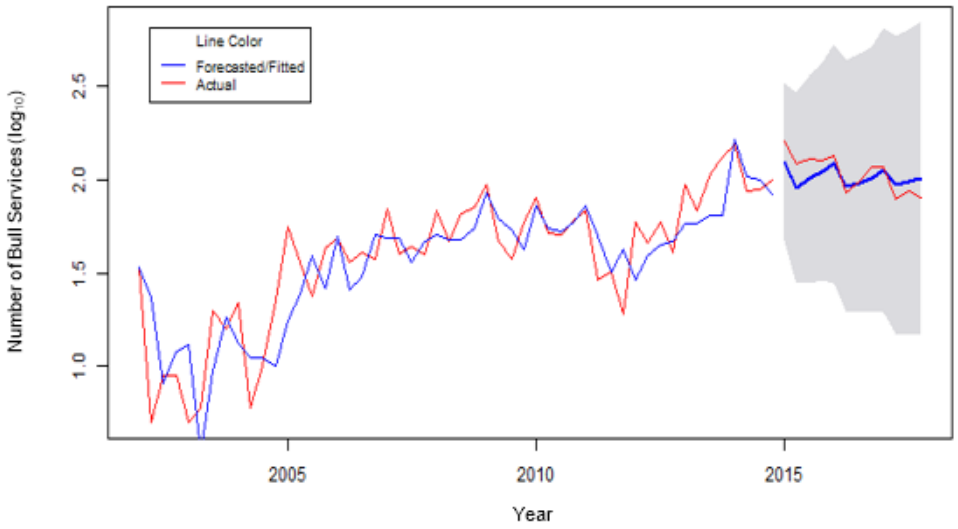


Figure 8: ARIMA 12-step ahead forecast

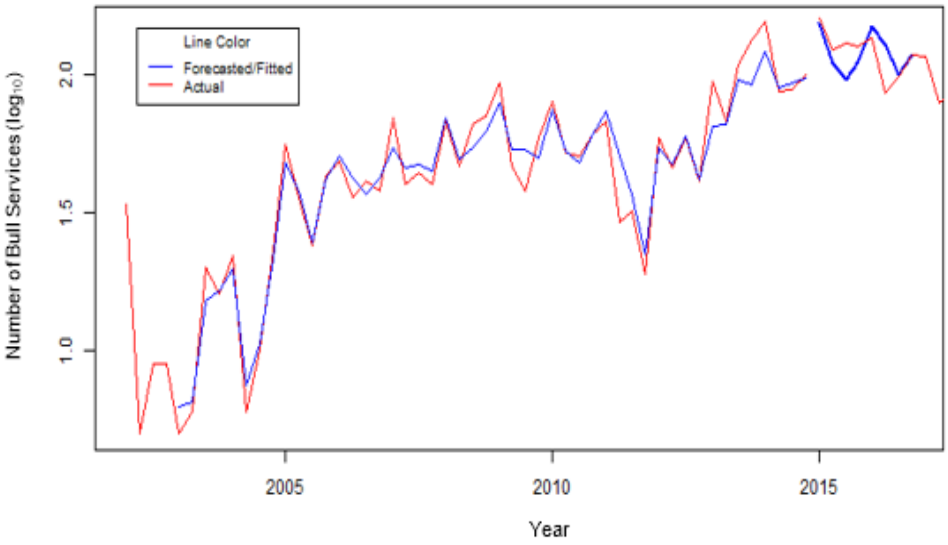


Figure 9: ANN 12-step ahead forecast.

Table 5: ARIMA and ANN model error (4-step ahead forecast).

Method	RMSE	MAPE (%)	MAE
ARIMA	0.13	6.45	0.12
Neural Network	0.02	0.55	0.01

Table 6: ARIMA and ANN model error (12-step ahead forecast).

Method	RMSE	MAPE (%)	MAE
ARIMA	0.08	3.31	0.07
Neural Network	0.07	2.77	0.06

4. Conclusion and Future Work

As a pioneering study in forecasting bull services, the experimental outcomes attained with data on the execution of ARIMA and ANN models in time sequence forecasting have been exhibited. The performance of the ARIMA and ANN predictive models established in this study was assessed. Though both models are showing good forecasting capability based on RMSE, MAPE, and MAE values, the ANN surpasses the traditional ARIMA in both 4-step and 12-step ahead bull natural mating forecasts. The ANN with a single hidden layer indeed can satisfactorily estimate any continuous function that agrees with theoretical corroborations. The ANN forecasting can surely help in optimizing the PCC-VSU's resources. A prognostic modeling of bull services using multivariate series is recommended.

Conflict of Interest

The authors declare that there is no conflict of interest in laying this paper to publication.

Acknowledgment

The authors' profound gratitude to the following individuals/ institutions: Roopam Upadhyay, Jiashen Liu, Visayas State University, Department of Computer Science and Technology-VSU, Philippine Carabao Center-VSU, Prof. Winston M. Tabada, Dr. Norberto E. Milla, and the College of Computer Studies of the Cebu Institute of Technology-University.

References

[1] J. B. Rola, C. L. C. Sta. Romana, L. S. Feliscuzo and I. F. M. Lopez, "ARIMA prognostic application to bull services for resource usage optimization," in *2019 International Conference on Multimedia Analysis and Pattern Recognition (MAPR)*, Ho Chi Minh, 2019, https://doi.org/10.1109/MAPR.2019.8743539.

[2] PCC, "Philippine Carabao Center annual report 2018," PCC, Muñoz, 2018.

[3] T. Hirata, T. Kuremoto, M. Obayashi, S. Mabu and K. Kobayashi, "Time series prediction using DBN and ARIMA," in *International Conference on Computer Application Technologies*, Matsue, Japan, 2016, https://doi.org/0.1109/CCATS.2015.15.

[4] B. Siregar and Fahmi, "Forecasting of raw material needed for plastic products based in income data using ARIMA method," in *5th International Conference on Electrical, Electronics and Information Engineering (ICEEIE)*, Malang, Indonesia, 2017, https://doi.org/10.1109/ICEEIE.2017.8328777.

[5] A. S. Katasev, L. Y. Emaletdinova and D. V. Kataseva, "Neural Network model for information security incident forecasting," in *International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, Moscow, Russia, 2018, https://doi.org/10.1109/ICIEAM.2018.8728734.

[6] S. Leholo, P. Owolawi and K. Akindeji, "Solar energy potential forecasting and optimization using Artificial Neural Network: South Africa Case Study," in *Amity International Conference on Artificial Intelligence (AICAI)*, Dubai, United Arab Emirates, 2019, https://doi.org/10.1109/AICAI.2019.8701372.

[7] J. C. Palomares-Salas, J. G. de la Rosa, J. G. Ramiro, J. Melgar, A. Aguera and A. Moreno, "ARIMA vs. Neural Networks for wind speed forecasting," in *International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA)*, Hongkong, 2009, https://doi.org/10.1109/CIMSA.2009.5069932.

[8] S. L. Ho, M. Xie and T. N. Goh, "A comparative study of Neural Network and Box-Jenkins ARIMA modeling in time series prediction," *Comput. Ind. Eng.,* vol. 42, pp. 371-375, 2002.

[9] M. K. Aiden and K. C. Tripathi, "A Neural Network model for time series forecasting of U.P. East Rainfall," in *International Conference on Advances in Computing, Communication Control and Networking (ICACCCN2018)*, Greater Noida (UP), India, India, 2018, https://doi.org/10.1109/ICACCCN.2018.8748734.

[10] M. Kumar and M. Anand, "An application of time series ARIMA forecasting model for predicting sugarcane production In India," *St. Bus. Econ.,* vol. 9, no. 1, pp. 81 - 94, April 2014.

[11] S. Makridakis, S. C. Wheelwright and R. J. Hyndman, Forecasting: methods and applications, New York, USA: John Wiley, 2008.

[12] S. A. DeLurgio, Forecasting principles and applications, Boston, MA, USA: Irwin McGraw-Hill, 1998.

[13] R. A. Yaffee and M. McGee, Introduction to time series analysis and forecasting: with applications of SAS and SPSS, 1st edition, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto: Academic Press, Inc., 2000.

[14] SAS Institute Inc., SAS/ETS®13.2 User's guide, Cary, NC, USA,: SAS Institute Inc., 2014.

[15] M. Akpinar and N. Yumusak, "Year ahead demand forecast of city natural gas using seasonal time series methods," *Energies,* vol. 9, no. 9, p. 727, 2016.

[16] A. Barnston, "Statistics how to," *Weather Forecast.,* vol. 34, no. 6, 1992, https://doi.org/10.1175/1520-0434(1992)007<0699:CATCRA>2.0.CO;2.

[17] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecasting,* vol. 22, no. 4, p. 679–688, 2006, hhttps://doi.org/10.1016/j.ijforecast.2006.03.001.

[18] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Clim. Res.,* vol. 30, pp. 79-82, 2005, https://doi.org/10.3354/cr030079.

[19] G. C. E. Bumanlag, M. P. Tolentino, E. D. Manlapig and E. P. Palacpac, "A case study on the implementation of the bull loan program in selected PCC regional centers," in *2nd International Conference on Agriculture, Environment and Biological Sciences (ICAEBS-15) August 16-17, 2015*, Bali, Indonesia, 2015.

[20] X. Li and Z. Zhang, "A time series forecasting based on ARIMA for Minqin value of output related to water resources," in *2012 International Symposium on Geomatics for Integrated Water Resource Management*, Lanzhou, 2012, https://doi.org/10.1109/GIWRM.2012.6349588.

[21] C. D. Lewis, "Industrial and business forecasting methods," *Butterworths,* vol. 2, pp. 194-196, 1982.

[22] A. A. Adebiyi, A. O. Adewumi and C. K. Ayo, "Comparison of ARIMA and Artificial Neural Networks models," *J. Appl. Math.,* vol. 2014, 2014.

[23] N. Kohzadi, M. S. Boyd, B. Kermanshahi and I. Kaastra , "A comparison of artificial neural network and time series models for forecasting commodity prices," *Neurocomputing,* vol. 10, pp. 169-181, 1996.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 423-429 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Alleviation of Nonlinear Impact Using PAPR Hybrid Technique in CO-OFDM Systems

Liqaa A. Al-Hashime[1,*], Sinan M. Abdul Satar[2], Ghaida A. Al-Suhail[1], Osama Saied[4]

[1]*University of Basrah, Department of Electrical Engineering, 964, Iraq.*

[2]*University of Technology, Department of Electrical and Electronic Engineering, 964, Iraq*

[3]*University of Gharyan, Department of Electrical and Electronic Engineering, 218, Libya*

ARTICLE INFO

Article history:
Received: 12 June, 2019
Accepted: 19 October, 2019
Online: 25 December, 2019

Keywords:
CO-OFDM
PAPR
BER
EVM
QF
CD
PMD
Nonlinear

ABSTRACT

Orthogonal Frequency Division Multiplexing (OFDM) is a modulation format that has recently attracted lots of consideration interior the long-haul fiber-optic transmission community. The most important advantage of optical OOFDM is its unlimited capability of canceling Inter-Symbol Interference (ISI) caused by Chromatic Dispersion (CD) and Polarization-Mode Dispersion (PMD). Specifically, Coherent OOFDM (CO-OFDM) presents a good and effective modulation method in modern optical communication systems. However, due to its high Peak to Average Power Ratio (PAPR), the performance of CO-OFDM is affected by nonlinear impairments. In this paper, we propose a new joint nonlinear technique; L3-by-3; with a distortion technique; clip-ping; that bypass the requirement for the use of any side information (SI) to minimize the high PAPR value of the transmitted signal and consequently improve various fiber nonlinear impacts through PAPR reduction. The simulation results reveal dependable and excellent signal recapture at the receiver and an effectively minimized PAPR level at the sender side.

1. Introduction

Because of the advantages and qualities of the OFDM system such as high spectral efficiency, this has led to its use in different modern communication applications. Furthermore, it is worthwhile noting that many authors have made the OFDM the first candidate for use in high-speed communications applications and high data rate systems. [1]. The OOFDM system is rated depending on the application in which it is used into two types, the first type is called Direct Detection OOFDM(DD), where a single photodiode is utilized for detection in the receiver.The second type is called Coherent Detection OOFDM (CO-OFDM) Which utilized an optical mixing is employed with a Local Oscillator (LO) in a coherent optical receiver. [2, 3]. The optical transmission system suffers from several disadvantages , especially when using a higher order of mapping M-PSK or M-QAM, where (M) is the order of mapping. these impairments have been discussed in many studies and articles and are classified as linear and non-linear. [4]. Nevertheless, coherent systems that utilizing Digital Signal Processing (DSPs) permits the compensation of system weakness. In addition, there are many ways to process the signal and compensate the impairments. Some of these processes are used in the receiving transmitter side, some of them employed in the optical channel, while the others are used in the receiver side.

The synchronization and balance method is an example of the methods used to compensate impairments in the optical channel as demonstrated in [5, 6]. Moreover, there are other methods that can be utilized to compensate the impairments that have occurred to the signal in the optical channel, such as Optical Phase Conjugate (OPC) to reduce nonlinearity, and Dispersion Compensated Fiber (DCF) to mitigate the impact of CD on the signal [7, 8]. Additionally, some methods that based on DSP are utilized in the transmitter side, such as phase-conjugated twin wave method [5], where some of such methods depend on diminishing the Peak to Average Power Ratio (PAPR) of the transmitted signal. A worthy reduction can also be gained when using one of these techniques, such as Selective Mapping (SLM) [6], Partial Transmit Sequence (PTS) [7], and non-linear transforms [8].

In this paper, the primary objective is to mitigate the impact of phase noise and nonlinearity in a CO-OFDM system employing 4QAM and 16QAM modulation through PAPR reduction in the system using the VPI software. The novel contribution of this article the use of the nonlinear method; L3by3; joined with the clipping distortion technique; for the first time in order to produce

*Corresponding Author: Liqaa A, Email: newliqas@gmail.com

a new combining technique to lessen the nonlinear impact through PAPR reduction for two modulation order with 10 Gbps bitrate.

2. CO-OFDM System Principle

In general, the concept behind OFDM is that one stream with high data rate is converted into different streams with lower data rates. It has been evince that these streams can be sent in parallel, offering more extensive symbol periods than single bearer style. Moreover, in OFDM these parallel streams are modulated by using one type of modulation such as, QAM, BPSK, and QPSK. [9]. The main factor behind the use of OFDM In modern radio norms is the high spectral efficiency that provides due to the presence of subcarriers which are orthogonal to each other. Because of its normal flexibility to reduce the impacts of chromatic dispersion [2]. Although the CO_OFDM system It is complex in design from the transmitter side but provides spectral efficiency, future sensitivity and immunity to polarization dispersion. CO_OFDM system divides into five sections: Radio Frequency (RF) OFDM transmitter, optical transmitter, optical channel, optical receiver and RF receiver. Figure 1 shows the primary sections of a CO-OFDM system.

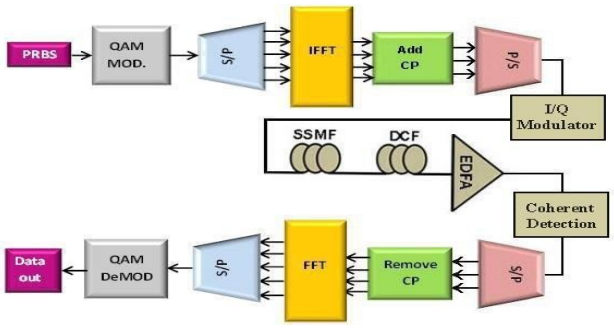


Figure 1 CO-OFDM system Block Diagram

In the transmitter, a serial binary data flow is obtained from a Pseudo-Random Binary Sequence generator (PRBS) and mapped into symbols using QAM mapping. In this study 4 and 16-QAM orders are considered where every two adjusted bits are mapped into one symbol in the case of 4-QAM, while in 16-QAM each symbol contains 4 bits. The serial data stream is converted to a parallel form before being passed to the Inverse Fast Fourier Transform (IFFT) block. In IFFT the block of data is transformed from the frequency domain into the time domain, and a cyclic prefix (CP) is added at the beginning of each time-domain OFDM symbol to eliminate Inter-Carrier Interference (ICI) and Inter-Symbol Interference (ISI) at the receiver. Afterward, the signal is converted from a parallel to a serial form (P/S) to send through the optical channel. The optical part of the system consists of an IQ modulator block which contains two Mach-Zehnder Modulators (MZMs) with a continuous wave (CW) laser that are utilized to up-converter the real/imaginary components of the signal, from the radio frequency(RF) domain to the optical domain.

In order to obtain the optical modulated signal, the signals that come from the two branches which have a phase difference of 90 are combined [10]. The subsequent part of the system is the fiber optic channel, which contains three processes: (i) Standard Single Mode Fiber (SMF) using to send optical signal.; (ii) compensating the dispersion in the optical channel using a Dispersion Compensating Fiber (DCF), where, the DCF, which is treated as a special kind of fiber optic that has a considerable negative dispersion, is used after the standard SMF [11] and (iii) amplifying the strength of signals being carried through the fiber optic channel using an Erbium-Doped Fiber Amplifier (EDFA) as well as compensating the attenuation induced by the optical fiber. The third part of the system is the receiver, which contains a coherent reveal block with laser used as a local oscillator, which is used to convert the optical signal to an RF signal. The coherent receiver can convert an optical-signal to a baseband electrical signal, where the 90° phase difference between the I and Q components is generated by optical hybrids and producing a phase shift equals of 180° among balanced detection. The final part of the system is the OFDM decoder, which consists of a serial to parallel converter, where data is converted from a serial into a parallel form, removing CP and executing a Fast Fourier Transform (FFT). To retrieve the transmitted information, the FFT is first used to turn the signal into the frequency domain. The outgoing data from the FFT is then switched to a sequential order to perform the de-modulation process and obtained the original data. The OFDM signal is involve of a substantial number of subcarriers. the addition of the power of (N) subcarriers that have the same phase will give top power (N) by the normal power. However, a huge PAPR increases the Nonlinear (FNL) impairments, which diminish the performance of the optical system. the OFDM consist of (N) data sequences of vector U(U_k (k=0,1,.., N-1)), which will be sent in side by side. One OFDM data symbol is:

$$u(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} U_k e^{2j\pi f_k t} \qquad (1)$$

where f_k is the subcarriers frequency spacing, U_k is the k_{th} information symbol vector and $u(t)$ is the baseband time domain OFDM signal after IFFT. the PAPR is relies upon the input information. As such, PAPR can be computed by finding the ratio between the peak power to the average power of the signal [12]. The equation of PAPR can then be written as:

$$PAPR = \frac{\max [|u(t)|\hat{}2]}{E[|u(t)|\hat{}2]} \qquad (2)$$

where $E[|()|^2]$ is the average power of the OFDM. The Cumulative Distribution Function (CDF) is the most frequently utilized parameters, and is utilized to gauge the performance of any PAPR technique. Regularly, the complementary cumulative distribution function (CCDF) is used when PAPR value exceeds the threshold. To find the probability that PAPR of an OFDM signal exceeds the threshold (p_0), assume the following complementary cumulative distribution function (CCDF) for non- overlapping sampling;

$$CCDF = p(PAPR > p_0) = 1 - (1 - \exp(-p_0))^N \qquad (3)$$

3. Proposed Technique

Generally, the nonlinear method L3-by-3 was first discussed in [14] for wireless OFDM system. This method is based on the Sliding Norm Transform (SNT) based PAPR reduction method [15]. Here, we propose, for the first time, using a combination of clipping and L3by3 in CO-OFDM system. Let x be a real vector with N samples, $x=(x1,x2,x3,\ldots,x_N)$.The sliding norm transformer that can be used in the transmitter side is then defined as follows:

$$y_n = \frac{x_n}{\sqrt[3]{\alpha + x_{n-1}^3 + x_n^3 + x_{n+1}^3}} \qquad (4)$$

The parameter (α) adjusts the PAPR of the transformed output. In other words, different values for PAPR can be obtained by changing α value, which indicates the possibility of setting a value

of PAPR by selecting the optimum value of α. The L3by3 transformer technique effectively lowers the PAPR of the CO-OFDM signals when compared with other techniques. The process of restoring x signals samples at the receiving end is done using the following Equation:

$$X = (+\sqrt[3]{x_n^3}\ sign\ (y_n) \quad n=1,2,3,.....N \qquad (5)$$

where $sign(0)$ is the Signum function. It can be noted that from Equation (5) data at the receiver can be recovered and the transmitted information can be obtained. The proposed hybrid technique is a combination between the clipping method and the L3-by-3 method and can be described by:

$$C(y_n) = \begin{cases} -A & if\ y_n<-A \\ y_n & if\ -A<y_n<A \\ A & if\ y_n>A \end{cases} \qquad (6)$$

where $C(yn)$ is the clipping signal, and A is a predetermined clipping level that can be defined as:

$$CR = \frac{A^2}{P_{in}} \qquad (7)$$

where P_{in} represents the mean input power of the OFDM signals, and CR is the clipping ratio. The principle of clipping technique is based on setting the maximum value of the input signal to a predetermined value if exceeded it will be clipped, otherwise, the input signal is transmitted without any change. In general, the clipping ratio is used to determine the level of clipping. More specifically, the clipping method is considered as a non-linear process that causes (in & out) band distortion. In other words, the out-band distortion causes spectral extension and can be taken away by filtering out the signal after clipping. On the other hand, the in-band distortion limits the Bit Error Rate (BER) and can be disposed of by performing clipping with the adequately oversampled OFDM signals (e.g., $L \geq 4$ where L is the oversample factor), resulting in a less degraded BER performance. To achieve an excellent PAPR performance, the L3by3 technique is combined with the clipping technique, where despite the increase in the BER of the system due to the clipping technique such an adverse of in-band distortion impact is eliminated, through the L3by3 technique, which reduces the number of signals exceeding the threshold value.

4. System Setup and Simulation Results

This section introduces the simulation of a CO-OFDM system with PAPR reduction using the VPI Transmission-Maker and MATLAB software packages. The global parameters set for the simulations and their values are presented in the Table (1).

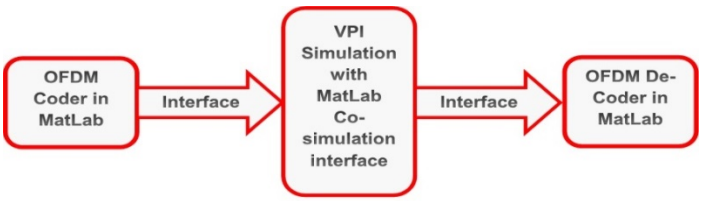


Figure 2 CO-OFDM system setup

Table1. The CO-OFDM system Simulation parameters

General parameters	
Bit-rate	10Gbps
Sample-rate	40Gbps
Sequence-length	8192
Samples-per-bit	2-bit
Mapping order	4QAM,16QAM
OFDM-parameters	
Sub-carriers	128
Cyclic-prefix	1/8
CW Laser & Local Oscillator	
Carrier freq.	193.1THz
Power in dBm	5dBm
Line-width	0.1MHz
SMF&DCF	
SMF length in km	50km
DCF length in km	5km
Number of loops	10
DCF attenuation	0.5dB/km
DCF dispersion	-160ps/nm/km
SMF dispersion	16ps/nm/km
SMF attenuation	0.2dB/km
Optical Amplifier DFA	
Gain	13dB
Noise Figure	4dB
α	0.2

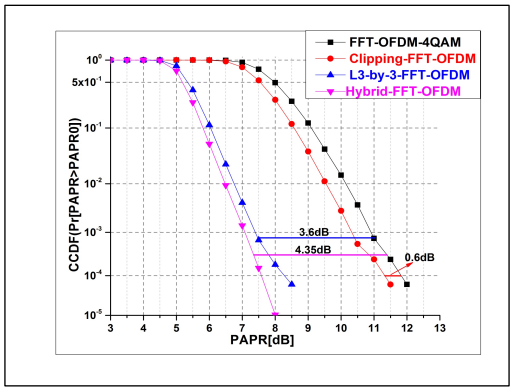
The OFDM signal generation and PAPR reduction were performed using MATLAB interfaced with VPI using the MATLAB Co-Simulation features, as shown in Figure (2).

First of all, referring to Equation (4) the control parameter α investigated for L3by3 for the 4-QAM and 16-QAM mapping orders. Table (2) shows the results obtained for PAPR at different values of α. The results show that α is directly proportional to PAPR. In other words, as α values change from 0 to 1, the value of PAPR increases, where the lowest PAPR value is obtained when the value of α is 0.1.

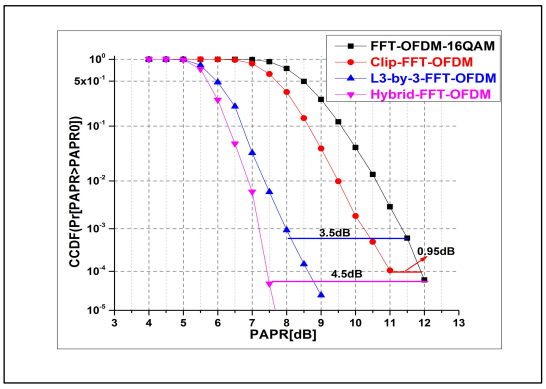
Table 2 PAPR values for different values of α

PAPR value in dB at probability of 10-4							
original	4QAM			11.5			
	16QAM			11.2			
α		0.1	0.2	0.4	0.6	0.8	1
L3by3	4QAM	6.35	7.75	8.4	9.1	9.4	10.1
	16QAM	6.75	7.6	8.5	9	9.55	10.2

Figure 3 (a) and (b) plot the CCDF as a function of PARP of the original signal without using any reduction technique for PAPR ; the CCDF when using a clipping for the signal at a clipping level 0.6, the CCDF when using L3by3 at α = 0.1; and the proposed hybrid technique for 4-QAM and 16-QAM optical OFDM signals, respectively.



(a) 4QAM



(b) 16QAM

(c)

Figure 3 A comparison of the CCDF vs. PAPR performance of a CO-OFDM system using a variety of techniques

Table 3 summarizes the results of PAPR reduction using the proposed technique at CCDF probability values of 10^{-4}.

Table 3 PAPR for the proposed technique at CCDF probability 10^{-4}

PAPR Value in dB at Probability 10-4				
	original	clipping	L3by3	L3by3+clip
4QAM	11.9	0.6	3.6	4.35
16QAM	11.9	0.95	3.5	4.5

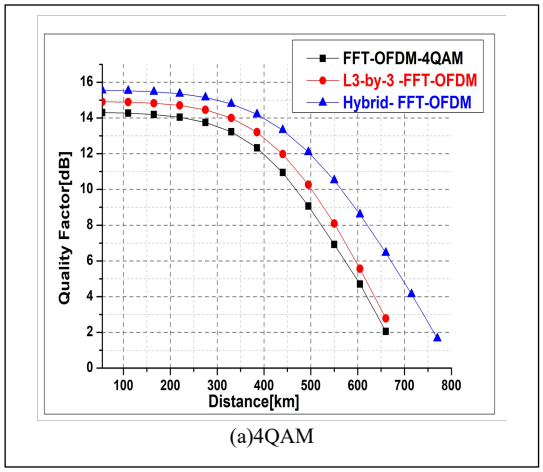
The table shows the efficiency of the proposed hybrid method in reducing the PAPR compared to using the L3by3 method and the clipping method separately. In addition, the suggested method has been compared with the other methods used to reduce PAPR in terms of their effectiveness in PAPR reduction as in table(4).

Table 4 Comparison between the suggested method and other PAPR reduction methods

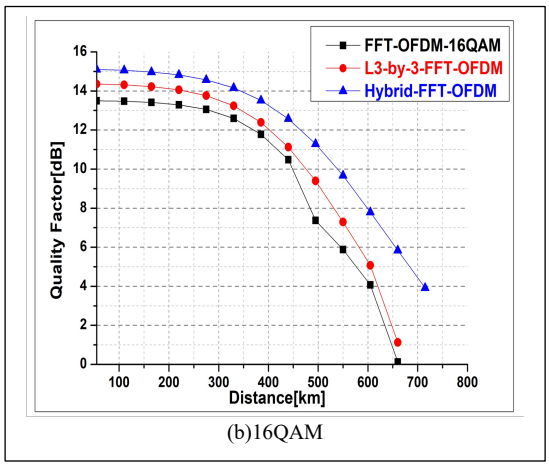
PAPR Value in dB at Probability 10-4				
Method	original	Hybrid	Reduction	Mapping
IPTS+CF[16]	10.5	7.32	3.18	QPSK
IPTS+MCF[16]	10.5	6.18	4.32	QPSK
Hadamard+m-law[17]	13.4	11.4	2	QPSK
L3by3+clipping	11.9	7.55	4.35	QAM

To check the efficiency and effects of the PAPR reduction on the system when changing distance, Figures (4) and (5) illustrate distance vs. Quality Factor (QF) and BER vs. distance respectively, for the system without using any technique to reduce PAPR ; for the system using the L3by3 ; and for the system using the proposed hybrid method. From figures, it is possible to deduce that the suggested hybrid method improves the QF and the BER for the system, where the results obtained for these parameters in the case of the hybrid method are better than those of the original system and those of system employing the L3by3 method. Table (5) summarizes the results obtained for QF from Figure (4).

As can be noted from Figure 4, the QF for the conventional CO-OFDM is good as well as the QF for both schemes at low distances (i.e. < 250 km), due to the reduction of the effect of chromatic dispersion as well as attenuation due to the presence of DCF that eliminates the effect of dispersion as well as the EDFA which minimizes the impact of attenuation. However, when we increase the launch power of the signal the distance will increases, thus increasing the nonlinear effects due to the Kerr effect, which reduces the performance of the system. As such for a larger distance (i.e. > 250 km), the proposed scheme outperforms the L3by3 method and the traditional CO-OFDM as it has the lowest PAPR, which reduces the fiber nonlinearity caused by the Kerr effect compared to the other two schemes (see Figure 3).



(a)4QAM



(b)16QAM

Figure 4 QF vs. distance for (a) 4-QAM and (b) 16-QAM

*Liqaa Abdul Sattar, Email:newliqas@gmail.com

Table.5. QF for CO-OFDM at different distances using different PAPR reduction techniques

Quality Factor in dB				
Distance [km]	Mapping	Original	L3by3	Hybrid
275	4QAM	13.7	14.4	15.2
	16QAM	13	13.7	14.5
550	4QAM	6.9	8	10.5
	16QAM	5.8	7.2	9.6

The 4-QAM and 16-QAM constellation points for the three aforementioned schemes presented in Table 5 show the QF improvement when the proposed scheme is considered compared to the other two schemes. For example, a QF improvement of 1.5 dB and 0.8 dB is obtained for the proposed scheme compared to the traditional CO-OFDM and L3by3 schemes, respectively, when 4-QAM is considered as a modulation order at a distance of 257 km. The table also illustrates that increasing the distance to 550 km for the same modulation order (i.e. 4-QAM) results in a 4.5 dB and 2.4 dB QF improvement for the proposed scheme compared to the traditional CO-OFDM and the L3by3 schemes, respectively.

Figure 5 shows the received constellation points of the traditional CO-OFDM, as well as the L3by3 and the proposed hybrid schemes. As illustrate in Figure 5 (a) for the 4-QAM mapping order, it is shown that an excellent constellation diagram for the original system (i.e. CO-OFDM) is obtained for the short fiber length with excellent QF. However, the 4-QAM constellation points of this scheme becomes worse for long fiber distances, as its QF becomes low. In the case of the L3by3 and the proposed Hybrid method, the performance remains good with increased distances, where a good quality factor can be obtained for up to 550 km for the system with the L3by3 method, and for up to 660 km for the system with the proposed Hybrid technique.

Additionally, Figure 5 (b) shows the constellation points of 16-QAM with and without the use of the L3by3 method and the proposed Hybrid method. The figure shows that the constellation points of the traditional CO-OFDM system is excellent for low fiber distances, but with increasing the distance, the signal power must be increased, which means increasing the nonlinearity and degradation of the system's quality factor. However, if PAPR reduction methods are used, the negative effect of the nonlinearity on the system performance is reduced and consequently the quality factor is improved, thus allowing increasing the transmission distance to 55 km in case of using the L3by3 method, and up to 100 km in case of using the proposed Hybrid method.

Figure (6) shows the relationship between BER and distance for the 4-QAM and 16-QAM systems. For the system with 4-QAM mapping order, a distance increase of more than 120 km and 95km is achieved for the proposed Hybrid and the L3by3 method, respectively, compared to the original system. As for the system with the 16-QAM mapping order, a distance increase of more than 105 km and 65 km is achieved for the suggested Hybrid and the L3by3 method, respectively, compared to the original system.

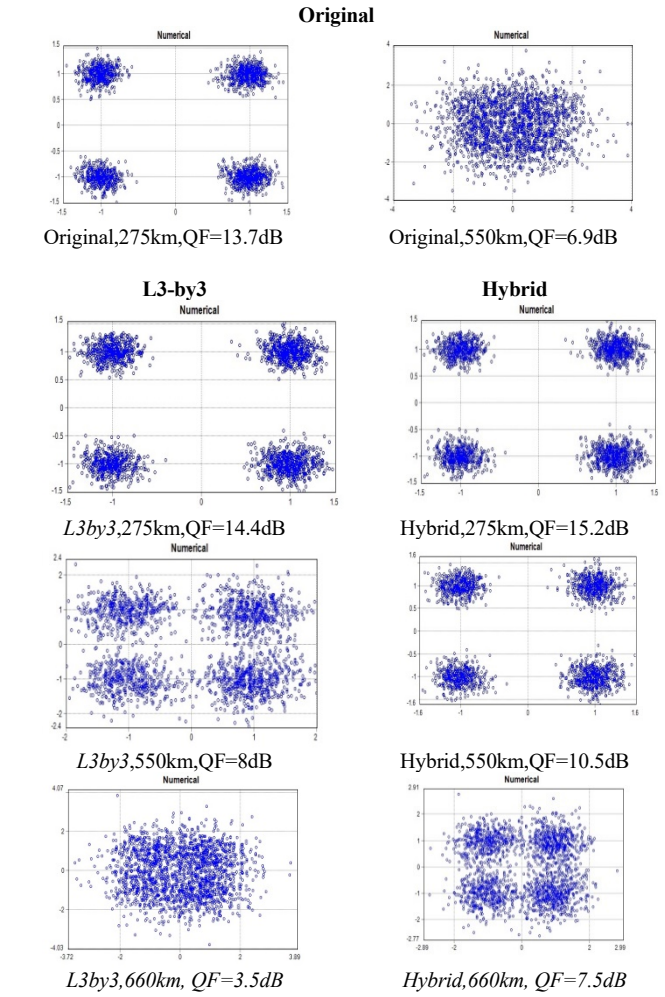


Figure 5 (a) 4-QAM Constellation diagrams of traditional CO-OFDM, L3by3and the proposed hybrid schemes at Different distance for 4-QAM

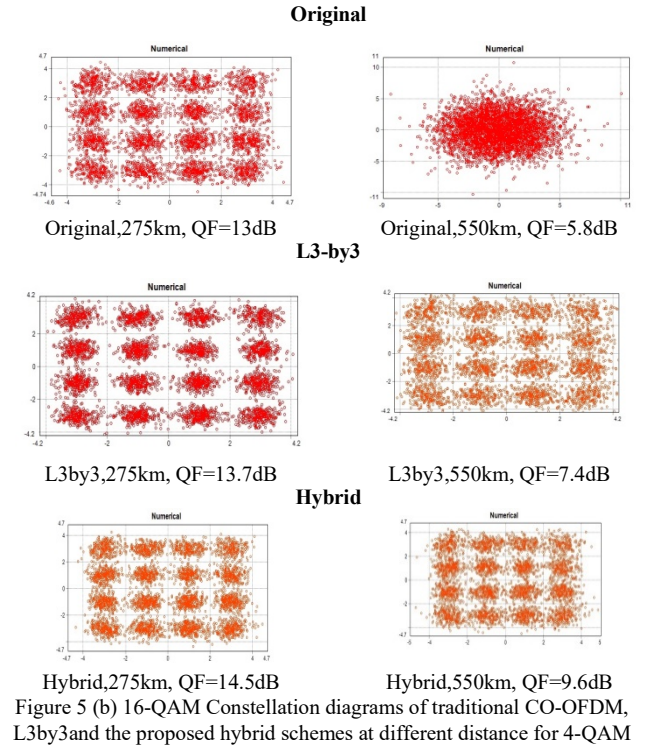
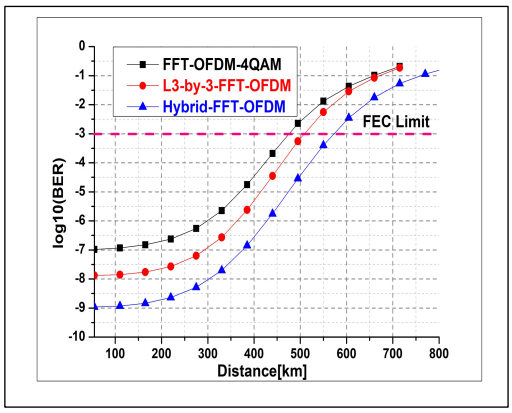
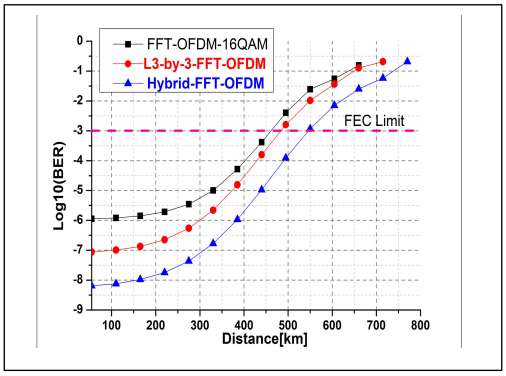


Figure 5 (b) 16-QAM Constellation diagrams of traditional CO-OFDM, L3by3and the proposed hybrid schemes at different distance for 4-QAM

Additionally, to examine the efficiency of the suggested Hybrid technique, the received power (RxP) and BER for the system with and without PAPR reduction techniques has been simulated. Figure (7) shows the results for the traditional system, the system with the L3by3 technique and the system with the proposed Hybrid technique at a distance of 550 km and launched laser power of 5 dBm. As shown in the figure, the best receiver sensitivity is obtained for the system with the Hybrid technique. Moreover, the figure demonstrates that the eye- opening of the system with the suggested combination technique is superior to the eye-opening in the case of using the L3by3 method or in the case of the primary system.



(a) 4 QAM



(b) 16 QAM

Figure 6 BER vs. distance

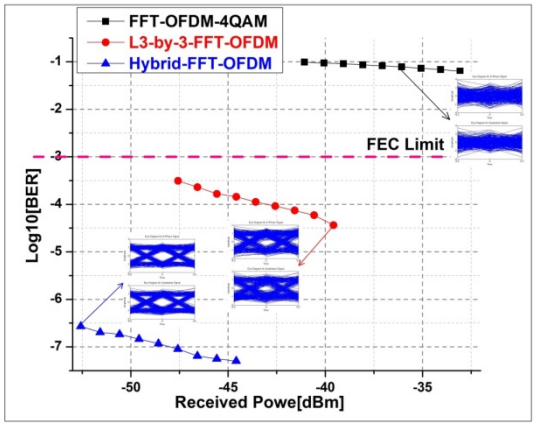
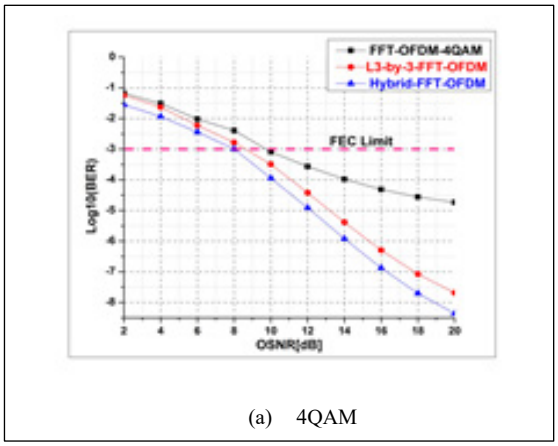
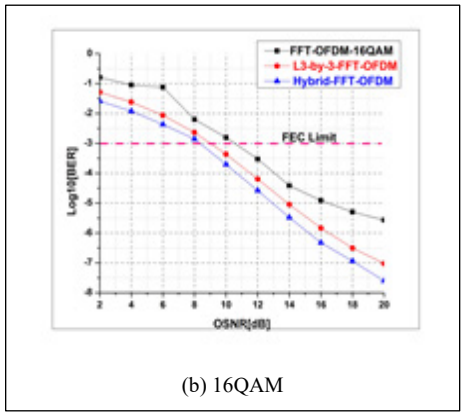


Figure 7 BER vs. received power at 550km for 4-QAM mapping order

Finally, Figure (8) shows the BER vs. OSNR for the traditional system, the system with the L3by3 method, and the system with suggested Hybrid reduction technique for PAPR at a distance of 550km and for 4-QAM and 16-QAM modulation orders. The results obtained show that when the value of the OSNR increases, the BER decreases.



(a) 4QAM



(b) 16QAM

Figure 8 OSNR vs. BER for (a) 4-QAM-(left), (b) 16-QAM-(right) at 550km

Table (6) provides a performance summary based on the BER vs. OSNR results for the 16-QAM and 4-QAMTable.

Table.6 Received OSNR for 4-QAM and 16-QAM at a distance of 550km and BER of 10^{-3}

OSNR in dB at 550km			
	original	L3by3	Hybrid
4 QAM	10	8.5	8
16 QAM	10.5	9	8.2

5. Conclusion

In this paper, an efficient and new hybrid method was proposed for PAPR reduction in CO-OFDM systems to lessen the impact of nonlinear effects and enhance the system performance. Both CO-OFDM with and without the proposed method were modeled using the VPI Transmission-Maker 9.5 and MATLAB software packages. The proposed scheme was shown to be an effective technique in reducing the PAPR. Furthermore, the scheme has low implementation complexity and no constraint on the system parameters because it does not need to send side information to the receiver. In addition, the results showed that the hybrid scheme can reduce the PAPR of the CO- OFDM by 4.35 dB and 4.5 dB for 4-

QAM and 16-QAM mapping orders, respectively. The hybrid scheme also achieved better QF and BER performance compared to the L3by3 and traditional schemes.

References

[1] Shieh, W. and C. Athaudage, Coherent optical orthogonal frequency division multiplexing. Electronics Letters, 2006, vol. 42, issue10, p. 587-589.

[2] Shieh, W., H. Bao, and Y. Tang, Coherent optical OFDM: theory and design. Optics Express, 2008. 16(2): p. 841-859.

[3] Schmidt, B.J., A.J. Lowery, and J. Armstrong, Experimental demonstrations of electronic dispersion compensation for long-haul transmission using direct-detection optical OFDM. Journal of Lightwave Technology, 2008. 26(1): p. 196-203.

[4] Nazarathy, M., et al. Recent advances in coherent optical OFDM high-speed transmission. In PhotonicsGlobal@ Singapore, 2008. IPGC IEEE.

[5] Le, S., et al. Phase-conjugated subcarrier coding for fiber nonlinearity mitigation in CO-OFDM transmission. In Optical Communication (ECOC), 2014 European Conference on, IEEE.

[6] Goebel, B., et al. PAPR reduction techniques for coherent optical OFDM transmission. In Transparent Optical Networks, ICTON'09. 11th International Conference on. 2009. IEEE.

[7] Dung, H.V.T., et al. PAPR reduction using PTS with low computational complexity in coherent optical OFDM systems. In IEEE 2012 18th Asia-Pacific Conf. on Communications.

[8] Rishi, P. and S. Tamilselvi, Mitigation of Non-Linear Effects Using Non-Linear Transform in Dispersion- Managed Coherent Optical OFDM Systems. Journal of Computational and Theoretical Nanoscience, 2018. 15(2): p. 551-557.

[9] Wang, Z., et al. Performance analysis of different modulation schemes for coherent optical OFDM system. In Communications and Mobile Computing (CMC), 2010 International Conference on, IEEE.

[10] Nehra, M. and D. Kedia, Design of Optical I/Q Modulator Using Dual-drive Mach-Zehnder Modulators in Coherent Optical-OFDM System. Journal of Optical Communications, 2018. 39(2): p. 155-159.

[11] Gnanagurunathan, G. and F.A. Rahman. Comparing FBG and DCF as a dispersion in the long haul narrowband WDM systems. In Wireless and Optical Communications Networks, 2006 IFIP International Conference on,IEEE.

[12] Kharagpur, B.K., U. Wali, and S. Bidwai. Novel technique to reduce PAPR in OFDM systems by clipping and filtering. In Advances in Computing, Communications, and Informatics (ICACCI), International Conference on. 2013. IEEE.

[13] Chen, L., et al. PAPR reduction in optical OFDM systems using asymmetrically clipping and signal scrambling technique. In International Conference on Optical Instruments and Technology: Optoelectronic Devices and Optical Signal Processing. 2015. International Society for Optics and Photonics.

[14] Palanivelan, S.A.M., Modified Sliding Norm Transform based approach for PAPR optimization in OFDM Systems. Australian Journal of Basic and Applied Sciences, 2013: p. p. 10.

[15] Grigoryan, A.M., S. Dursun, and M.M. Grigoryan. Method of sliding norm transforms for peak-to-average power reduction in OFDM systems. In Industrial Electronics, 2006 IEEE International Symposium on. IEEE.

[16] Tong, Z.-r., Y.-n. Hu, and W.-h. Zhang, PAPR reduction in CO-OFDM systems using IPTS and modified clipping and filtering. Optoelectronics Letters, 2018. 14(3): p. 209-211.

[17] Xiao, J. et al., Hadamard transform combined with a companding transform technique for PAPR reduction in an optical direct-detection OFDM system. Journal of optical communications and networking, 2012. 4(10): p. 709-714.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 430-442 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Linear Logic Synthesis of Multi-Valued Sequential Circuits

Nikolay Butyrlagin[1,*], Nikolay Chernov[1], Nikolay Prokopenko[1,2], Vladislav Yugai[1]

[1]*Information System and Radio Engineering Department, Don State Technical University, Gagarin Sq. 1, Rostov-on-Don, 344000, Russia*

[2]*Institute for Design Problems in Microelectronics of Russian Academy of Sciences, IPPM RAS, Zelenograd, 124681, Russia*

ARTICLE INFO

Article history:
Received: 02 October, 2019
Accepted: 10 December, 2019
Online: 25 December, 2019

Keywords:
Memory element
Logic operation
Linear representation
Truncated difference operation
Comparison operation
Circuitry implementation

ABSTRACT

The basics of unconventional design of current circuits of two-valued and multi-valued memory elements (ME) for storing current digital signals and flip-flops of the main types on their basis are considered. A nontraditional method of ME synthesis is proposed, which is based on the mathematical tool of linear algebra. Linear equations, structural and functional schemes of the main types of logic elements for the construction of ME are given. CMOS-circuitry of various versions of realization of linear logic elements of multi-valued MEs is considered. The comparative analysis is carried out and advantages and disadvantages of linear realization of current two-valued memory elements and flip-flops on their basis in comparison with potential (Boolean) realization are defined. The forecast of merits and demerits of linear implementation of multi-valued memory elements is given.

1. Introduction

Nowadays a potential logic and its mathematical tool, that is Boolean algebra, dominate in the world of logic synthesis of products of microelectronics. We sweepingly accept this fact and appreciate its contribution to digital electronics development.

However, as long as Boolean algebra exists, there have been attempts to replace it by another mathematical tool. It results in efforts to apply spectral transformations of Krestenson, Walsh, Haar [1-5], polynomial representations [6-9] for logic synthesis.

The authors of this article suggest applying mathematical tool of linear algebra for the aims of logic synthesis [10]. As a result of investigations, we have defined primary application domain of this tool that is current circuits of random valuedness. The Boolean algebra is not convenient to synthesize multi-valued digital circuits, so linear algebra is not an opponent to the Boolean algebra despite the specific nature of the logic synthesis process and multi-valued digital circuits circuitry [11-19]; it is rather an addition to Boolean algebra .

The authors have developed:

- Mathematical tool and methods of linear logical synthesis of two-valued and multi-valued digital structures;
- circuit solutions of the basic logic elements for constructing digital structures;
- circuitry solutions of some digital combinational two-valued and multi-valued structures (adders, decoders, multiplexers, digital components of the ADC, etc.);

In the process of research, the computer simulation of two-valued and multi-valued logical elements proved their operational integrity [20 - 27]. The effectiveness of linear algebra as a mathematical tool of logical synthesis is confirmed by patents RU2506695, RU2509412, RU2514789, RU2547225, RU2553070, RU2549144 and publications [20 - 27].

The results of research show the advantages of the proposed design methodology over the Boolean: it enables to suggest logic elements with improved characteristics compared to Boolean elements, as well as workable multi-valued elements. The purpose of this article is to demonstrate the possibilities of using current logic (linear algebra) as a mathematical tool for logical synthesis and circuit design of sequential two-valued and multi-valued digital structures using current flip-flops as an example.

2. Methodology

At the beginning, the well-known structural schemes of potential flip-flops of the main, widely used types are given. Then, structural schemes of their two-valued current analogues are

*Corresponding Author: Nikolay Butyrlagin, Don State Technical University, Russia | Email: butyrlagin@gmail.com

synthesized. Further, the features of the structural completeness of multi-valued flip-flops and their circuitry implementation are considered.

Two-valued potential flip-flops. Two-valued flip-flops have a quit wide classification [11, 12]. RS-, D-, T- and JK-flip-flops are the most commonly used two-valued flip-flops. A brief description of the structural implementation of these flip-flops is given below.

Asynchronous RS-flip-flop is a tandem amplifier with positive feedback. Its structural circuit, implemented on AND-NOT (OR-NOT) elements is in Fig. 1 a, and its logic symbol is in Fig. 1 b.

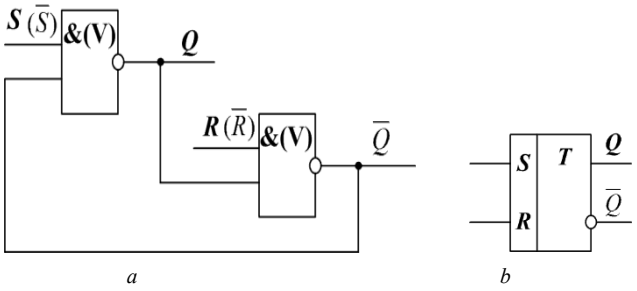


Figure 1: Structural circuit (*a*) and logic symbol (*b*) of *RS*-flip-flop on OR-NOT (AND-NOT) elements.

Its structural circuit, implemented on inhibit circuit ("implication gate") is shown in Fig. 2 *a*, its logic symbol is shown in Fig. 2 *b*.

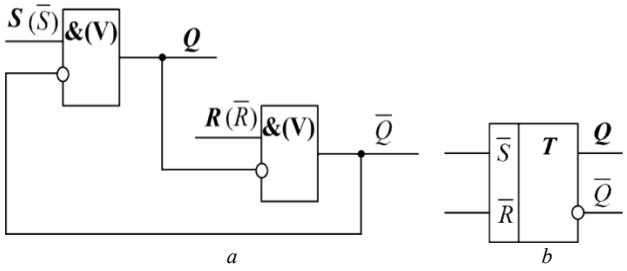


Figure 2: Structural circuit (*a*) and logic symbol (*b*) of *RS*-flip-flop on implication gates (inhibit circuits).

Single-ended synchronous RS-flip-flop is characterized by additional synchronization input. The flip-flop is switched to opposite state, if there is enabling level of signal at synchronization input. Its logic symbol is shown in Fig. 3.

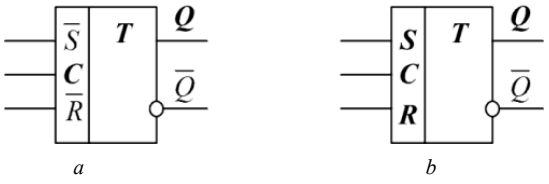


Figure 3: Logic symbol of single-ended synchronous *RS*-flip-flops: a – on OR elements, b – on AND elements

To eliminate disadvantages of single-ended *RS*-flip-flop's switch [12] the push-pull circuits are used. Their structures, implemented on AND-NOT (OR-NOT) elements are in Fig. 4 *a*, *b*.

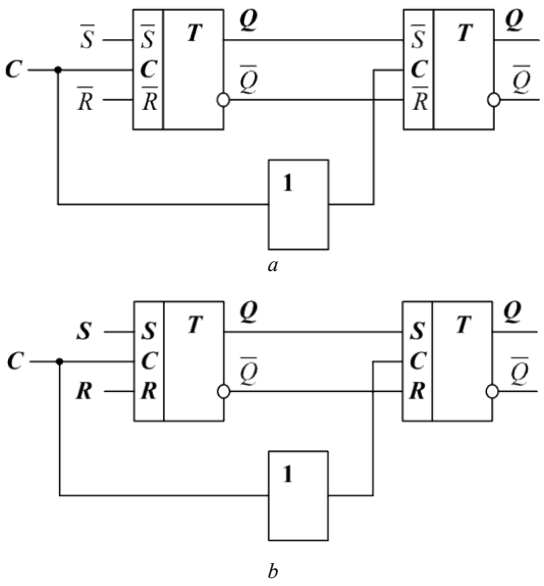


Figure 4: Structural circuits of push-pull RS-flip-flops.

The logic symbol of push-pull *RS*-flip-flops is in Fig. 5.

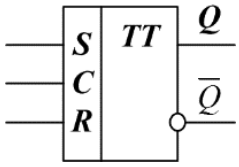


Figure 5: Logic symbol of push-pull RS-flip-flop.

Now let us show the structures of flip-flops of other popular types. The D-flip-flop structural circuit is shown in Fig. 6 a, its logic symbol is shown in Fig. 6 b.

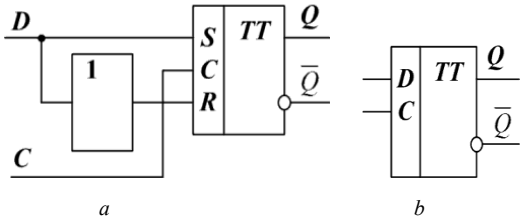


Figure 6: D-flip-flop: *a* – structural circuit, *b* – logic symbol.

The structural circuit of *T*-flip-flop is shown in Fig. 7 a and its logic symbol is shown in Fig. 7 b.

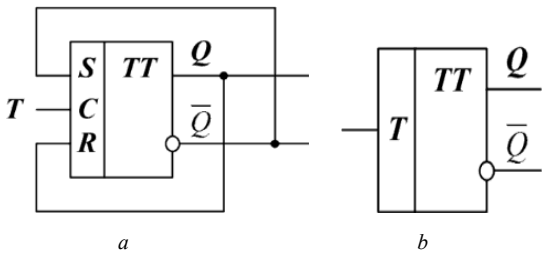


Figure 7: T-flip-flop: *a* – structural circuit, *b* – logic symbol.

And finally, the structural circuit of *JK*-flip-flop is in Fig. 8.

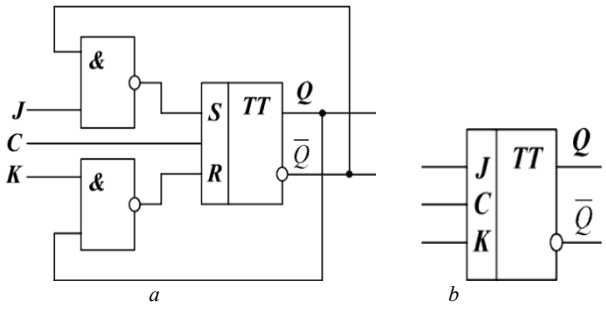


Figure 8: JK-flip-flop: *a* – structural circuit, *b* – logic symbol.

Two-valued current flip-flops. The potential flip-flops are similar to current flip-flops at system technology level, so it is possible to show the current asynchronous structural circuit of *RS*-flip-flop in the form, shown in Fig. 9, and its logic symbol - in the form, specified in Fig. 10 [21].

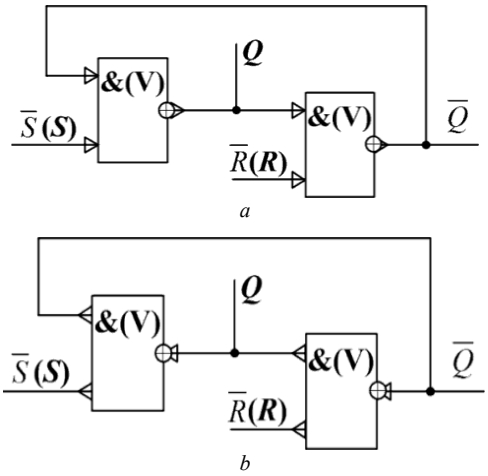


Figure 9: Structural circuit of two-valued asynchronous RS-flip-flop on AND-NOT (OR-NOT) current elements: *a* – output implementation, *b* – input implementation.

In the structural and functional diagrams given hereinafter, arrows indicate the directions of the current input and output signals: incoming arrows indicate input currents, outcoming arrows indicate output currents. The type of implementation of the element (incoming or outcoming) is determined by the direction of the output signal of the element.

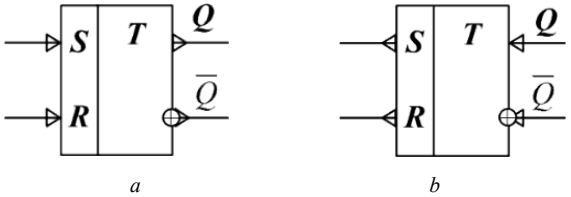


Figure 10: Logic symbol of two-valued asynchronous RS-flip-flop on AND-NOT (OR-NOT) current elements: *a* –output implementation, *b* – input implementation.

Possible differences in the options for implementing flip-flop circuits appear at the levels of functional and circuit design.

The structural circuit of synchronous *RS*-flip-flop is in Fig. 11 and its logic symbol is in Fig. 12.

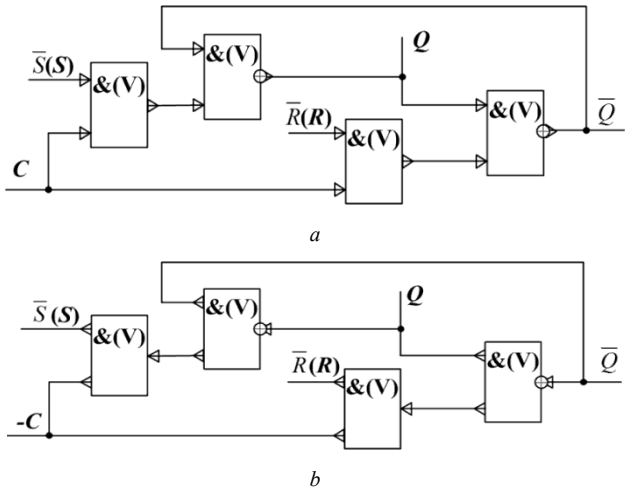


Figure 11: Structural circuit of two-valued synchronous RS-flip-flop on AND-NOT (OR-NOT) current elements: *a* –output implementation, *b* – input implementation.

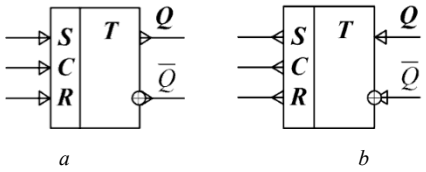


Figure 12: Logic symbol of two-valued asynchronous RS-flip-flop on AND-NOT (OR-NOT) current elements: *a* – output implementation, *b* – input implementation.

Similarly, structural implementation of push-pull current synchronous *RS*-flip-flop is similar to corresponding potential circuit. The structures of synchronous push-pull RS-flip-flops are in Fig. 13 and their logic symbols - in Fig. 14. The "+" element performs the function of an inverter.

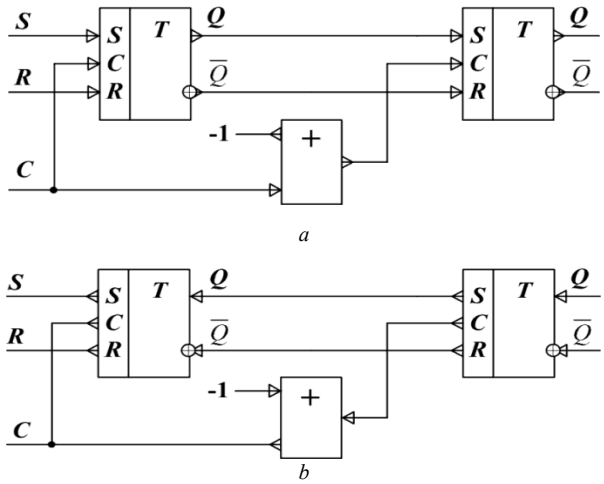


Figure 13: Structural circuits of push-pull synchronous current RS-flip-flops: *a* – output implementation, *b* – input implementation.

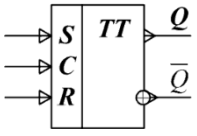


Figure 14: Logic symbol of push-pull synchronous current RS-flip-flops

The structural implementations for flip-flops of other types are also similar to implementations of corresponding potential circuits.

3. Three-valued flip-flops

Multi-valued flop-flops. Usually the multi-valued digital circuits are considered as generalization of corresponding two-valued circuits [1]. But bivalent logic degeneracy causes problems for such generalization. The primary difficulties of multi-valued flip-flop synthesis are the following:

– selection of logic valuedness, which is suitable to construct multi-valued memory elements;

– selection of multi-valued equivalents of logic operations, applied for logic synthesis of memory elements in two-valued logic;

- multi-valued interpretation of inversion.

Let us consider the above problems in details.

As to select logic valuedness to construct multi-valued flip-flops we should note the following. One uses concept of structural completeness (transition and outputs system completeness) for flip-flops, applied as automata memory elements for functional completeness of logic functions (basis) for combinational circuit in automata theory [9]. By default, the flip-flop is characterized by transition and output system completeness, if it is possible to assign some set of control signals to each output state, the said signals transit it from any previous state to the said state within finite number of steps. In two-valued case any known memory elements has this property.

The transition graph of two-valued flop-flop can be represented in the form shown in Fig. 15.

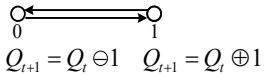


Figure 15: Transition graph of two-valued flop-flop.

In the multi-valued case, the situation is somewhat more complicated.

When k = 3 (Fig. 16), the multi-valued flip-flop can have transition graphs of direct (per neighboring vertex: 0→1→2→0 and through vertex: 0→2→1→0) and reverse (per neighboring vertex: 0→2→1→0 and through vertex: 0→1→2→0) type, which are opposite to each other. They are also complete graphs, covering all vertices, i.e. three-valued flip-flop has transition and output system completeness.

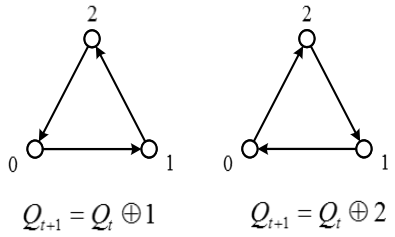


Figure 16: Direct and reverse transition graph for three-valued flip-flop.

When k = 4 (Figure 17), the multi-valued flip-flop can also have direct and reserve graphs.

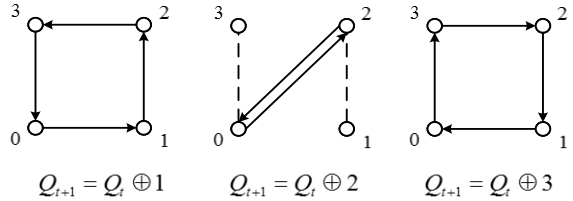


Figure 17: Four-valued memory element transition graphs.

However, from Fig. 17 it follows that the four-valued flip-flop does not have the completeness of the system of transitions and outputs: when passing through two vertices, the transition graph is looped. The transition graphs for k = 5 and k = 7 are in Fig. 18 and Fig. 19.

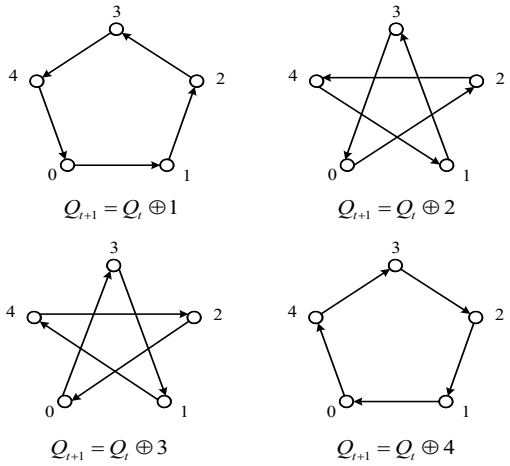


Figure 18: Transition graph of five-valued flip-flop.

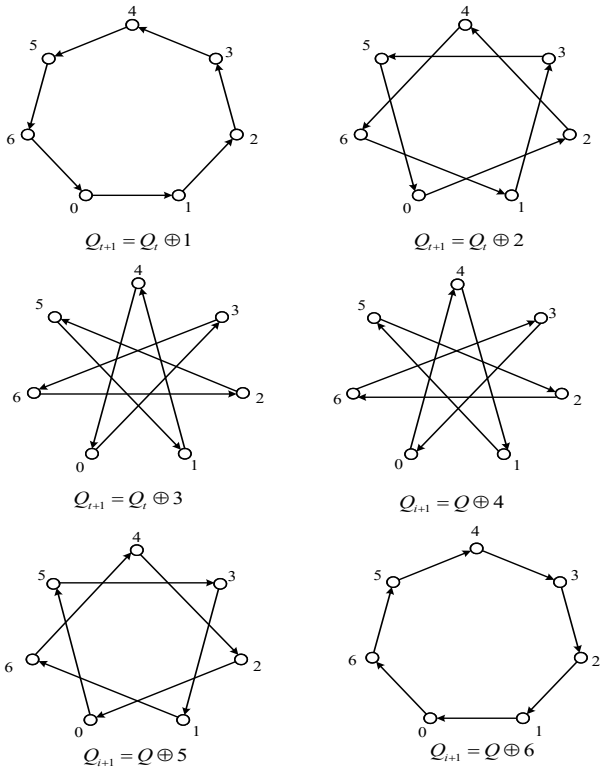


Figure 19: Seven-valued flip-flop transit graph.

The Fig. 15 – 19 show a confirmation of the known fact [9], that the transition and output system completeness of the flip-flop of random valuedness is provided, only when logic valuedness magnitude is a prime number.

Let us note the following, concerning selection of multi-valued equivalents of two-valued operations, used for memory elements synthesis. The most popular flip-flops for two-valued version are flip-flops, implemented on the elements realizing operations $\overline{\&}(x_1, x_2)$ (AND-NOT) and $\overline{\vee}(x_1, x_2)$ (OR-NOT).

The following operations are functionally close to these operations: $\overline{\&}(x_1, \bar{x}_2)$ ("Inhibit per x_2 ") and $\overline{\vee}(x_1, \bar{x}_2)$ ("Implication from x_1 to x_2"). Generalization of these operations for multi-valued version results in the following representations:

$$\overline{\&}(x_1, x_2) \rightarrow \begin{cases} \min(x_1, x_2) \oplus i \\ \min(x_1, x_2) \ominus i \end{cases};$$

$$\overline{\vee}(x_1, x_2) \rightarrow \begin{cases} \max(x_1, x_2) \oplus i \\ \max(x_1, x_2) \ominus i \end{cases};$$

$$\overline{\&}(x_1, \bar{x}_2) \rightarrow \begin{cases} \min(x_1, x_2 \oplus i) \\ \min(x_1, x_2 \ominus i) \end{cases};$$

$$\overline{\vee}(x_1, \bar{x}_2) \rightarrow \begin{cases} \max(x_1, x_2 \oplus i) \\ \max(x_1, x_2 \ominus i) \end{cases}.$$

In current circuits the authors use elements, implementing comparison operations $>(x_1, x_2))$, sum module $\oplus(x_1, x_2)$ and difference $\ominus x_1 x_2$ and truncated difference operations.

$$x_1 \dot{-} x_2 = \begin{cases} x_1 - x_2 \text{ when } x_1 \geq x_2 \\ 0 \text{ when } x_1 < x_2 \end{cases}.$$

If to combine it with linear space operation, it is possible to have a variety of different presentations for multi-valued equivalents for main two-valued functionally complete system AND, OR, NOT. There are some possible versions of these operations in Tables 1-4.

Table 1: Primary operations, presented through truncated difference

Boolean representation	Linear difference representation
$\min(x_1, x_2)$	$x_1 \dot{-} (x_1 \dot{-} x_2)$
$\max(x_1, x_2)$	$x_1 + (x_2 \dot{-} x_1)$
$x_1 \oplus x_2$	$x_1 + x_2 - k[1 \dot{-} (k \dot{-} (x_1 + x_2))]$
$x_1 \ominus x_2$	$x_2 - x_1 + k\{1 \dot{-} [k \dot{-} (x_2 \dot{-} x_1)]\}$
$x \oplus i$	$x + i - k[1 \dot{-} (k \dot{-} (x + i))]$
$x \ominus i$	$i - x + k[1 \dot{-} (k \dot{-} (i \dot{-} x))]$

It is possible to synthesize more than 80 different flip-flops, using only set of operations, specified in the above tables, which are required to form logic elements, from which flip-flop memory elements are constructed.

It is necessary to note, that being fully equivalent to AND, OR, NOT operations in two-valued version, in multi-valued case some of these expressions results in functions, different from multi-valued equivalents of AND, OR, NOT operations. This confirms the two-valued logic degeneracy, because multi-valued functional systems, different from traditional nature, are converged to it.

Table 2: Primary operations, presented through sum and difference module

Boolean two-valued representation	Linear-module representation
$\min_3^{M2}(x_1, x_2)$	$\frac{x_1 + x_2 - \lvert x_1 - x_2 \rvert}{2}$
$\max_3^{M2}(x_1, x_2)$	$\frac{x_1 + x_2 + \lvert x_1 - x_2 \rvert}{2}$
$\oplus_3^{M2}(x_1, x_2)$	$\lvert x_1 + x_2 \rvert$
$\ominus_3^{M2}(x_1, x_2)$	$\lvert x_1 - [(k-1) - x_2] \rvert$
$\oplus_3^{M2}(x, i)$	$\lvert x + i \rvert$
$\ominus_3^{M2}(x, i)$	$\lvert x - [(k-1) - i] \rvert$

Table 3: Primary operations, presented through comparison operation

Boolean representation	Linear-compared representation
$\min_3^{c2}(x_1, x_2)$	$x_1 - (x_1 > x_2)$
$\max_3^{c2}(x_1, x_2)$	$x_2 + (x_1 > x_2)$
$\oplus_3^{c2}(x_1, x_2)$	$x_2 - x_1 + k(x_1 > x_2)$
$\ominus_3^{c2}(x_1, x_2)$	$x_1 - x_2 + k(x_2 > x_1)$
$\oplus_3^{c2}(x, i)$	$i - x + k(x > i)$
$\ominus_3^{c2}(x, i)$	$x - i + k(i > x)$

Table 4: Primary operations, presented by threshold form

Boolean representation	Linear-threshold representation
$\min_3^{c2}(x_1, x_2)$	$[(x_1 > 0 + x_2 > 0) > 1] +$ $+[(x_1 > 1 + x_2 > 1) > 1]$
$\max_3^{c2}(x_1, x_2)$	$x_1 + x_2 - [(x_1 > 0 + x_2 > 0) > 1] -$ $-[(x_1 > 1 + x_2 > 1) > 1]$
$\oplus_3^{c2}(x_1, x_2)$	$x_1 + x_2 + 3[(x_1 > 1 + x_2 > 1) > 1] -$ $-3[(x_1 > 0 + x_2 > 1) > 1] -$ $-3[(x_1 > 1 + x_2 > 0) > 1]$
$\ominus_3^{c2}(x_1, x_2)$	$x_1 + x_2 - 2[(x_1 > 0 + x_2 > 0) > 1] -$ $-3[(x_1 > 1 + x_2 > 1) > 1] +$ $+[(x_1 > 1 + x_2 > 0) > 1]$
$\oplus_3^{c2}(x, i)$	$x_1 + i + 3[(x_1 > 1 + i > 1) > 1] -$ $-3[(x_1 > 0 + i > 1) > 1] -$ $-3[(x_1 > 1 + i > 0) > 1]$
$\ominus_3^{c2}(x, i)$	$x_1 + i - 2[(x_1 > 0 + i > 0) > 1] -$ $-3[(x_1 > 1 + i > 1) > 1] +$ $+[(x_1 > 1 + i > 0) > 1]$

Finally, on the multi-valued interpretation of inversion. In the two-valued case, the flip-flop consists of two identical logic elements, therefore, you can submit information to the input and "remove" it from the flip-flop output in a direct or inverse code. By controlling the process of recording information in a flip-flop, it is possible to process information by setting a direct or inverse value at the desired output.

A multi-valued flip-flop has k inputs and the same number of outputs. Here, the information processing function is much "richer" than in the two-digit case: controlling the process of recording information in such a flip-flop, it can be cyclically shifted left or right ("rotate"). In the process of information rotating, you must remember that

$$x \oplus i = x \ominus (k - i).$$

This ratio enables, if necessary, to replace one of the operations of the recorded ratio with another.

The synthesis of multivalued flip-flops should also be preceded by an agreement on the principles of generalizing two-valued flip-flops to a multivalued case. The agreement proposed by the authors of this article consists of the following main provisions:

- the number of flip-flop information inputs must be equal to the logic valuedness;

- at the outputs of the flip-flops there must be a full set of logic values (the number of outputs of the flip-flop must be equal to the valuedness of the logic);

- the valuedness of the output signals of the functional elements of the flip-flop should not exceed the valuedness of the logic used.

The listed provisions, as is easy to see, in the two-valued case are fully implemented.

RS flip-flop. A verbal description of the functioning algorithm of a multi-valued RS-flip-flop corresponding to the agreement on the principles of generalizing two-valued flip-flops to a multi-valued case is given below:

- the storage mode corresponds to the values of the input signals S = SR = R = k - 1 (or S = SR = R = 0);

- the signal S increases the flip-flop state index relative to the current state towards the state k-1, the SR signal decreases or increases the current one depending on the ratio of the current state index and the signal value at this input, and the signal R decreases towards the state 0; change of state in a cycle is impossible (let's leave it for universal flip-flops!);

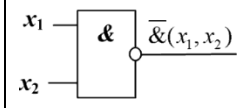
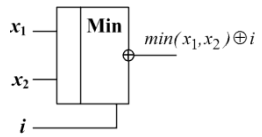
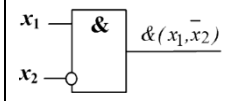
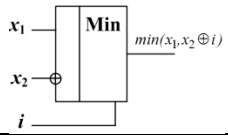
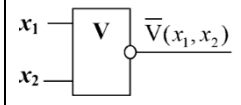
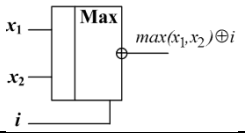
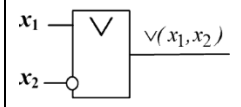
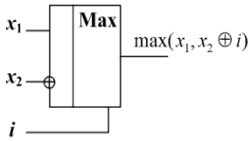
- the magnitude of the change in the state index is equivalent to the value of the input signal: a signal equivalent to 1 can increase (or decrease) the index of the current state by 1, a signal equivalent to 2 by 2, etc .;

- to change the flip-flop state index relative to the current state i, signal S can take values from 1 to k-1-i (larger values will be equivalent to k-1-i in action), similarly, signal R can take values from 1 to i- 1 (larger values will be equivalent to i -1);

- Allowed combinations of input signal values are all combinations in which one of the input signals is k - 1 (0) and the others take any of the possible values.

The meaning of the last paragraph of the algorithm is that to change the state of the flip-flop, the two inputs must have signals 0 or k-1, which passes the setting signal of any of the allowed values, and the third - the signal that sets the state.

Table 5: Logic symbols for "traditional" multi-valued equivalents of two-valued operations

Operation	Logic symbol	
	Two-valued	Multi-valued
$\min(x_1, x_2) \oplus i$		
$\min(x_1, x_2 \oplus i)$		
$\max(x_1, x_2) \oplus i$		
$\max(x_1, x_2 \oplus i)$		

4. Results and Discussion

Structural design. The above research results allow us to move to the structural design of current two-valued and multi-valued flip-flops.

In the following exposition, we will mainly confine ourselves to the synthesis of flip-flops with the valuedness k = 2, 3, since reasoning about flip-flops of greater valuedness practically repeats similar arguments about ternary flip-flops. In necessary cases, the features of flip-flops of greater valuedness will be specially discussed.

To construct the structural, functional and basic circuit diagrams of current flip-flops, the above current operations and their combinations, as well as linear space operations, will be used. Moreover, we restrict ourselves to the "traditional" approach, i.e. using operations generalized to the multi-valued case in the "conventional" way.

The logic symbols for "traditional" multi-valued equivalents of above operations are in Table 5.

A base of current operations is the following $[\min(\max)](x_1, x_2)$ and $[\oplus\ (\ominus)](x, i)$operations with different representations, using various current operations. The most important representations (see Tables 1-4) are specified below:

– operation $\min(x_1, x_2)$:

$$\min(x_1, x_2) \Rightarrow \begin{cases} x_1 \dot{-} (x_1 \dot{-} x_2) \\ \frac{x_1 + x_2 - |x_1 \dot{-} x_2|}{2} \\ [(x_1 + x_2) > 3] + [(2x_1 + x_2) > 2] \end{cases};$$

– operation $\max(x_1, x_2)$:

$$\max(x_1, x_2) \Rightarrow$$

$$\begin{cases} x_1 + (x_2 \dot{-} x_2) \\ \frac{x_1 + x_2 + |x_1 \dot{-} x_2|}{2} \\ x_1 + x_2 - [(x_1 + x_2) > 3] - [(x_1 + 2x_2) > 2] \end{cases};$$

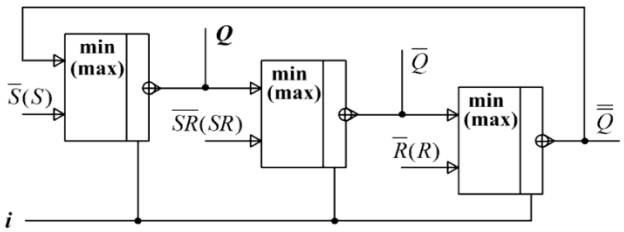
– operation $\oplus\,(x, i)$:

$$\oplus\,(x, i) \Rightarrow \begin{cases} i - x + 3(x > i) \\ x + i - k[1 \dot{-} (k \dot{-} (x + i))]; \\ |x + i| \end{cases}$$

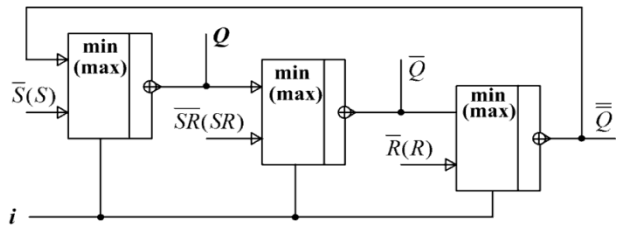
– operation $\ominus\,(x, i)$:

$$\ominus\,(x, i) \Rightarrow \begin{cases} |x - [(k - 1) - i]| \\ x - i + 3(i > x) \\ |x - i| \end{cases}.$$

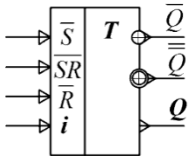
The logic elements, from which the analogues of conventional flip-flops are constructed, are consecutive connection of elements implementing $[\min(\max)](x_1,x_2)$ and $[\oplus\,(\ominus)](x, i)$ operations.



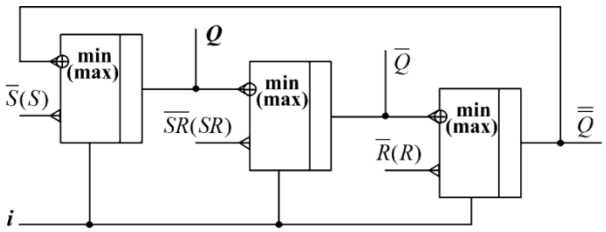
a



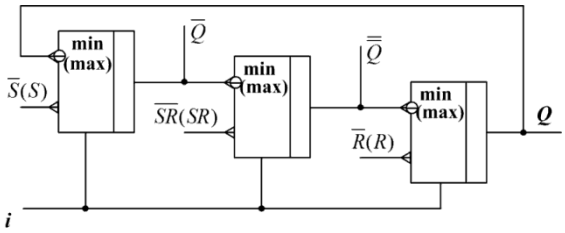
b



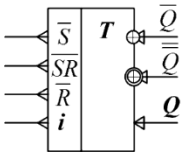
c



a



b



c

The structural circuit and logic symbol of three-valued asynchonous RS-flip-flop on equivalent elements AND-NOT (OR-NOT) are in Fig. 20-21.

Let us consider functioning of three-valued RS-flip-flop, which structural circuit is shown in Fig. 20 a. It is seen that state storage phase is provided by signals of level 2 (min – realization) or 0 (max – realization) at inputs S, SR, and R.

Let in *min* implementations of the RS-flip-flop (Fig. 20), at the beginning, the setup signal *i* is 0, which corresponds to a direct (i.e., without rotation) recording of the value of the input information. When the input signal is set S = 0 - 2 and SR = R = 2, the signal $\overline{Q} = S \oplus 1$is formed at the output of the logic element D2, at the next (D3) - signal $\overline{\overline{Q}} = \overline{Q} \oplus 1$, at the last (D1) - signal $Q = \overline{\overline{Q}} \oplus 1$. For example, at S = 0, the signals $\overline{Q} = 1$, $\overline{\overline{Q}} = 2$, $Q = 0$ (state "0") are generated at the outputs.

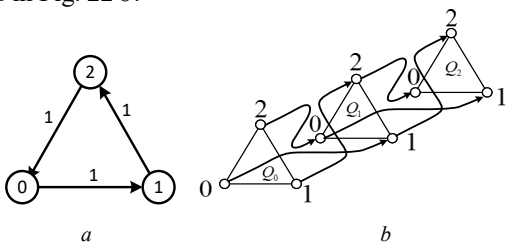
If the setting signal *i* is 1, which corresponds to the right cyclic shift of the signal by one in each logic element, then the signal $S \oplus 2$ is generated at the output of the first logic element D2, – $\overline{\overline{Q}} = \overline{Q} \oplus 2$ at the output of the second element (D3), the output of the third (D1) is the signal $Q = \overline{\overline{Q}} \oplus 2$. For example, with S = 0 and SR = R = 2, the signals $\overline{Q} = 2$, $\overline{\overline{Q}} = 1$, $Q = 0$ (state "0") are formed at the outputs.

If the setting signal *i* is 2, which corresponds to the left cyclic shift of the signal by two (to the left cyclic shift by one) in each logic element, then the signal $S \oplus 3 = S$ is generated at the output of the first logic element (D2), – $\overline{\overline{Q}} = \overline{Q} \oplus 3 = \overline{\overline{Q}}$ at the output of the second element (D3), the signal $Q = \overline{\overline{Q}} \oplus 3 = Q$ at the

output of the third (D1). For example, with S = 0 and SR = R = 2, the signals $\overline{Q} = 0$, $\overline{\overline{Q}} = 1$, $Q = 2$ (state "2") are formed at the outputs.

Let the flip-flop be in state "1". In this case, the input signal SR can transit it into any of the remaining states, since this signal determines the value of the output signal of the element D3 and propagates with a shift along the remaining elements of the flip-flop, the input state of the flip-flop cannot be changed, since the second input of the element D3 is 0, and the signal S can transit the flip-flop only in the state "0".

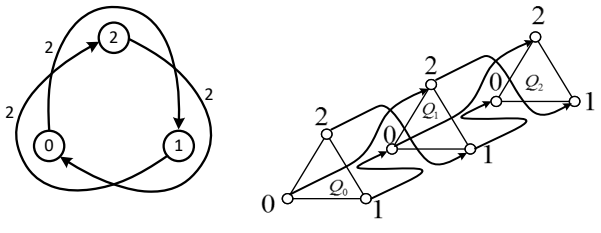
The graph of changes of the state, when i = 1 is in Fig. 22 *a* and the process of signal values change on the outputs of logic elements is in Fig. 22 *b*.



a *b*

Figure 22: Three-valued RS-flip-flop graph, when *i* = 1: *a* – state change, *b* – output signal value change.

When i = 2, a switch process will go reverse. The graphs of three-valued RS-flip-flop for such case are in Fig. 23.



a *b*

Figure 23: Three-valued RS-flip-flop graphs, when *i* = 2: *a* – state change, *b* – output signal value change.

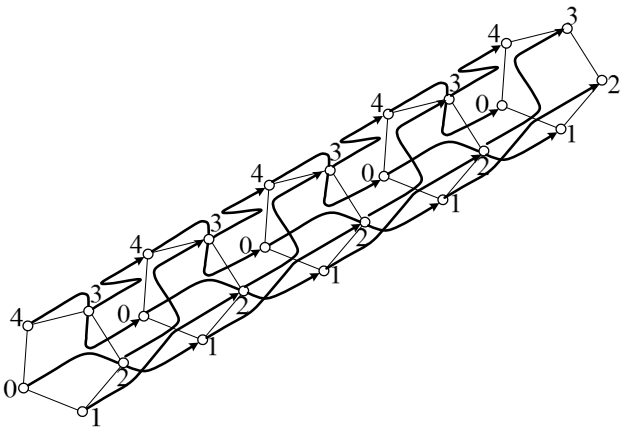


Figure 24: Graph of the value changes of the output signals of five-valued RS-flip-flop *i* = 1.

Similarly, it is possible to construct graphs for state change and values of flip-flop output signals of higher valuedness. The structure of the graph will change greatly at other values of k. The graph of value changes of the output signals of five-valued flip-plop at *i*=1 is in Fig. 24 as an example.

Here visualization is lower than the one for three-valued case.

As in two-valued case synchronous single-ended RS-flip flops are characterized by wider functional possibilities. Their main difference from asynchronous flip-flops is an availability of additional synchronization input *C*, a signal on which defines a flip-flop state change moment, if there is a signal, which is able to result in such change, at informational input (inputs). An input signal is able to be from 0 to *k*-1 at any input, so signal *C* should take on value 0 ("lock" input of flip-flop, its state is not changed at any input signal change) or *k*-1 (any value signal comes to flip-flop input and it goes to the state, corresponding to the said value). It is in good agreement with two-valued flip-flops, if we understand synchronous signal inversion as $1 - C$, that is $(k - 1) - C$.

It is possible to describe k-valued single-ended synchronous RS-flip-flop as RS-flip-flop, where S-, RS- and R-inputs are connected to S-, RS- and R-inputs of memory element through *min* or *max*, where the synchronous signal C is supplied to second inputs.

The RSC-flip-flop structural circuit is in Fig. 25 and its logic symbol is in Fig. 26.

Push-pull circuit of three-valued RSC-flip-flop is constructed from two single-ended RSC-flip-flops similar to two-valued case. Its structural circuit and logic symbol are in Fig. 27.

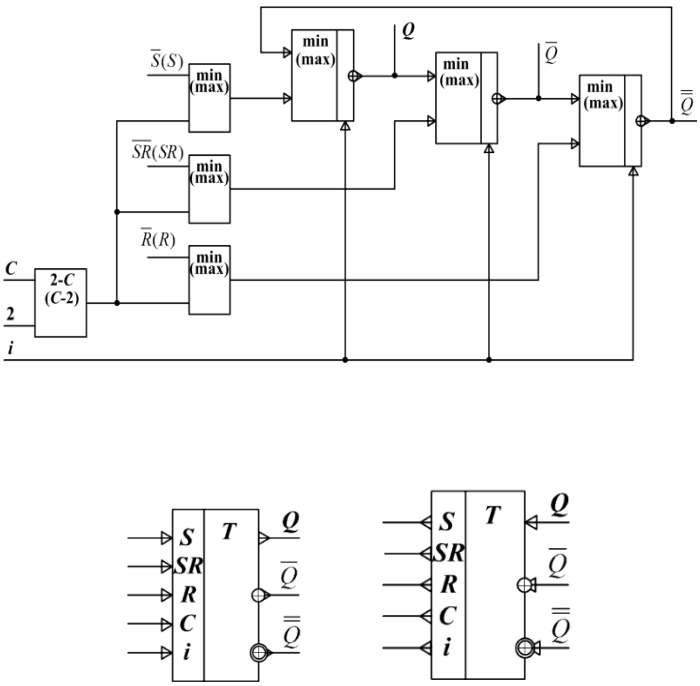


Figure 26: Logic symbol of three-valued synchronous RSC-flip-flops: *a* –output implementation, *b* – input implementation.

The tree-valued single-ended synchronous D-flip-flop is functionally similar to two-valued D-flip-flop: it is set to state, correspodning to input signal value by synchronous signal. In other words, it is possible to describe three-valued D-flip-flop as RS-flip-flop with input signals:

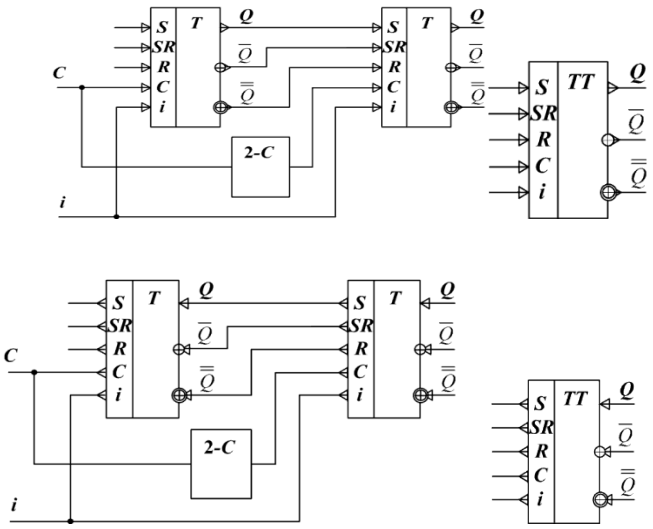


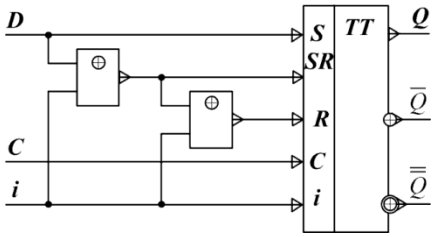
Figure 27: Structural circuit and logic symbol push-pull synchronous RS-flip-flop: *a* – output implementation, *b* – input implementation.

$$S = D;$$

$$RS = D \oplus i;$$

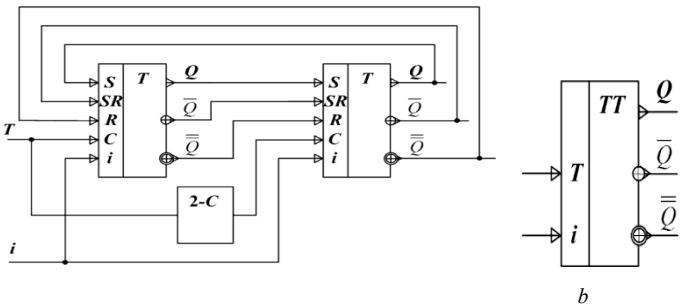
$$R = D \oplus i \oplus i.$$

The structural circuit of synchronous single-ended k-valued D-flip-flop (AND-NOT – equivalent) is in Fig. 28.



A structure of OR-NOT – equivalent of three-valued D-flip-flop is of similar form.

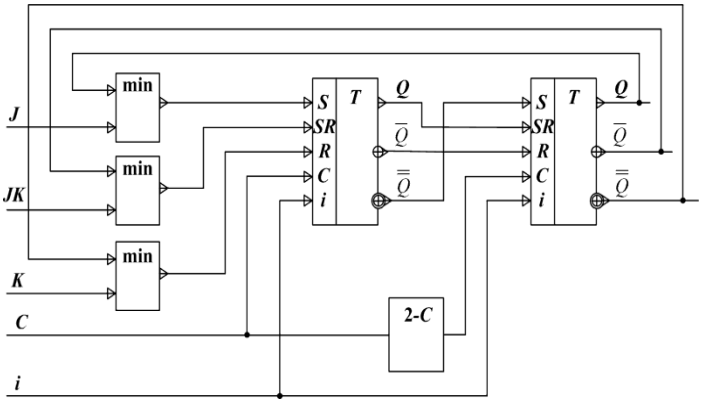
The three-valued T-Flip-Flop is also implemented on two single-ended RS-flip-flop and is counter module 3, an order of counting is defined by value of i. The structural circuit and logic symbol of T-flip-flop are in Fig. 29.



b

Figure 29: Three-valued push-pull synchronous T-flip-flop: *a* – structural circuit, *b* – logic symbol.

The structural circuit of synchronous three-valued push-pull JK-flip-flop is in Fig. 30.



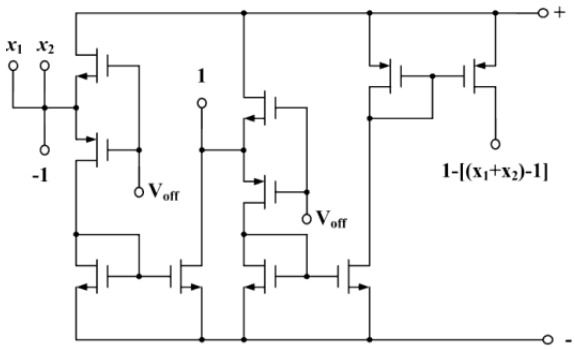
We perform circuit design via example of two-valued current flip-flops and their three-valued traditional generalization on the base of conjunctive function representation. Let us consider design of two-valued and three-valued logic elements and the simplest flip-flops. We do not consider circuits of multi-valued flip-flops of higher valuedness because it is quite difficult to show the said circuits within this article.

Two-valued flip-flops. All variants of basic operation representation, specified in Tables 1-4, are suitable for their synthesis.

The following operation is a current equivalent of OR-NOT operation in difference form:

$$1 \dot{-} [(x_1 + x_2) \dot{-} 1].$$

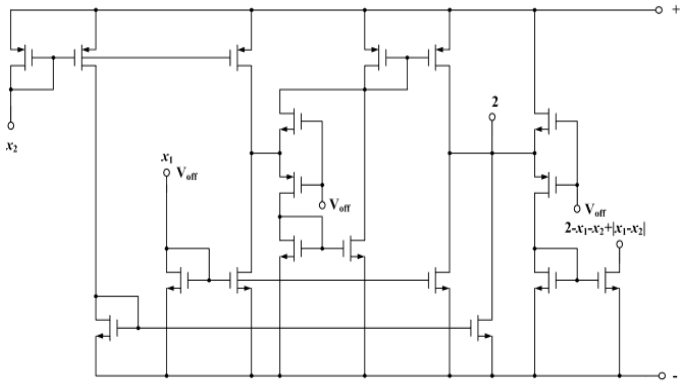
It is possible to verify its correctness by substitution. The basic diagram of logic element, which implements the said operation, is in Fig. 31.



Applying module form of current representation results in the following form of OR-NOT operation:

$$2 - x_1 - x_2 + |x_1 - x_2|.$$

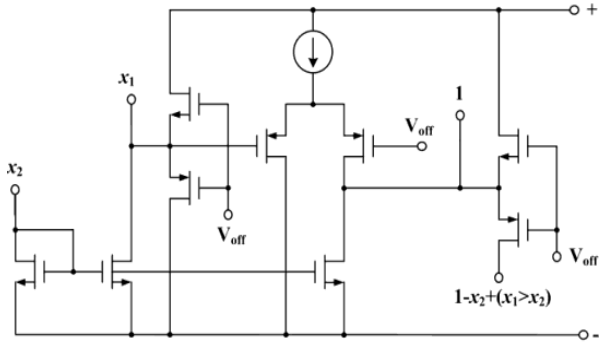
The basic diagram of OR-NOT operation implementation current variant in module form is in Fig. 32.



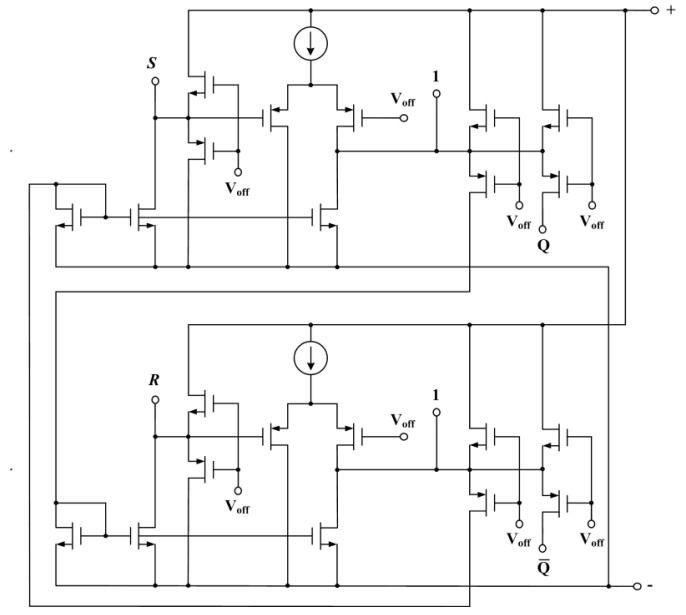
A comparison form of OR-NOT current representation results in the following form:

$$1 - x_2 + (x_1 > x_2).$$

Its circuit implementation is in Fig. 33.



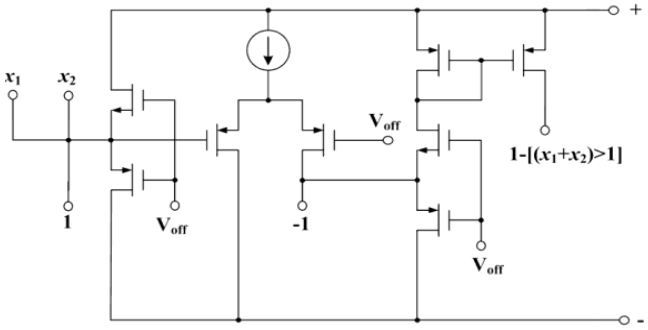
When combining two elements in amplifier circuit with positive feedback, it results in two-valued RS-flip-flop construction. Its basic diagram is in Fig. 34.



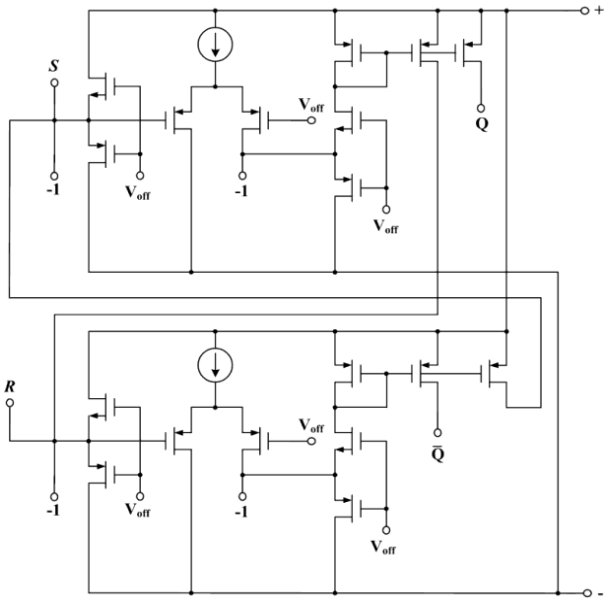
OR-NOT operation threshold current representation results as following:

$$1 - [(x_1 + x_2) > 1].$$

Its circuit implementation is shown in Fig. 35.



The simple current RS-flip-flop in threshold representation is in Fig. 36.



Multi-valued flip-flops are constructed on elements-generalizations of two-valued elements, which implement OR-NOT operation. As we have already said above, the traditional generalizations are consecutive connection of elements $[\min(\max)](x_1, x_2)$ and $[\oplus\ (\ominus)](x, i)$, that is $\min(x_1, x_2) \oplus i$, $\min(x_1, x_2 \oplus i)$, and two similar operations, based on $\ominus$. The operations, specified in Table 3, are not suitable for this purpose, because multiple generalizations are different from traditional ones here. It is better to consider them separately.

Difference representation of $\min(x_1, x_2) \oplus i$ is the following:

$$[x_1 \dot{-} (x_1 \dot{-} x_2)] + i - 3\langle 1 \dot{-} \{3 \dot{-} [[x_1 \dot{-} (x_1 \dot{-} x_2)] + i]\}\rangle$$

The implementation circuit for this expression is in Fig. 37.

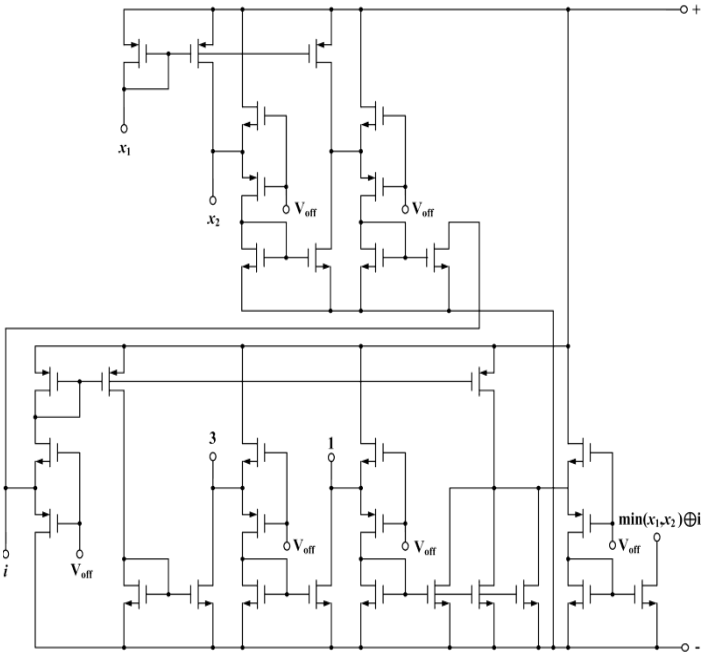
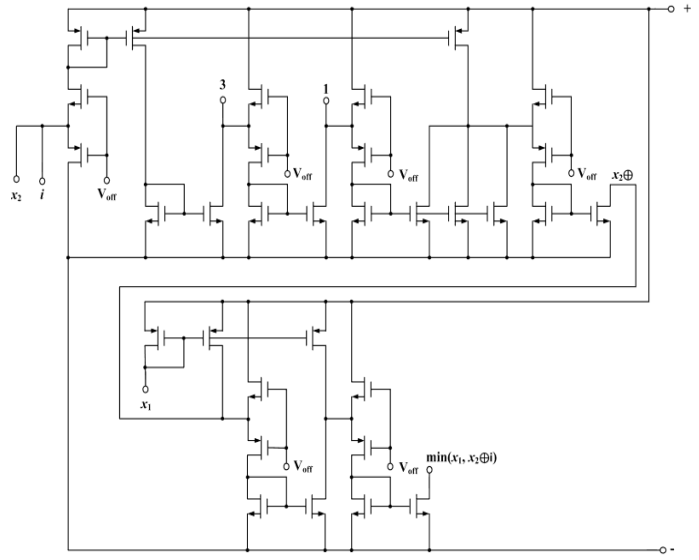


Figure 37: Basic diagram to implement difference representation $\min(x_1, x_2) \oplus i$.



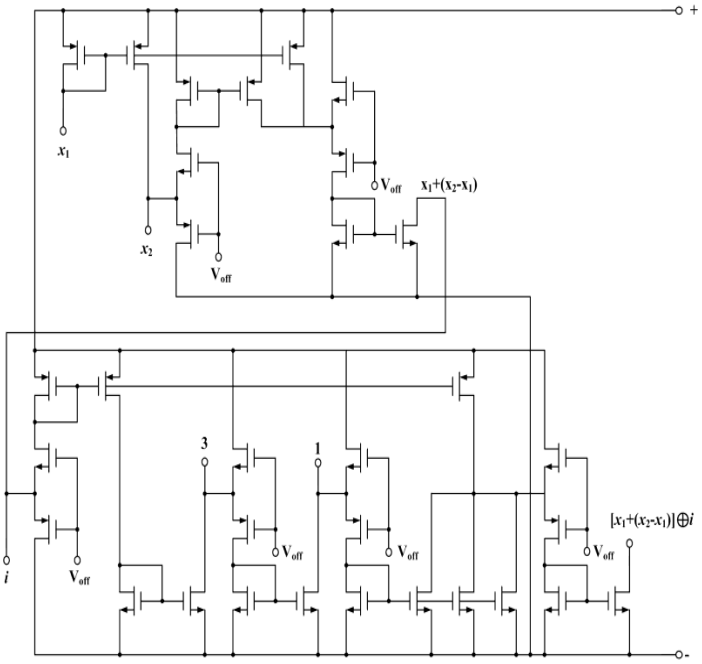
$x_1, x_2 \oplus i$).

A difference representation of $\min(x_1, x_2 \oplus i)$ is the following:

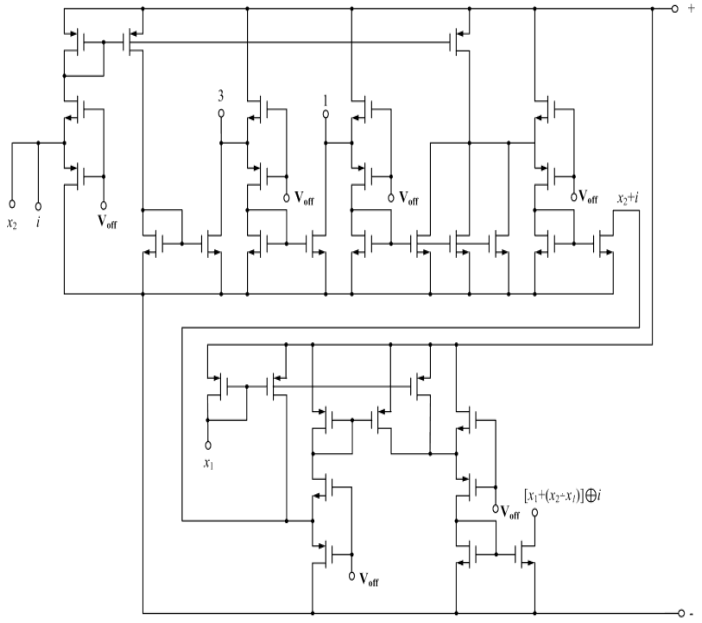
$$i - [x_1 \dot{-} (x_1 \dot{-} x_2)] +$$
$$+3\langle 1 \dot{-} \{3 \dot{-} [i \dot{-} [x_1 \dot{-} (x_1 \dot{-} x_2)]]\}\rangle$$

It is visible that the last circuit is different from the previous one only by order of element sequence of $\min(x_1, x_2)$ and $x \oplus i$ operation implementation.

The circuit implementation of $\max(x_1, x_2) \oplus i$ and $\max(x_1, x_2 \oplus i)$ logic elements is in Fig. 39.



a



b

a – $\max(x_1, x_2) \oplus i$, *b* – $\max(x_1, x_2 \oplus i)$.

A module representation of $\min(x_1, x_2) \oplus i$ is formed from $\min(x_1, x_2)$ representation as following:

$$\frac{x_1 + x_2 - |x_1 - x_2|}{2}$$

$\oplus i$ which has the following form:

$$3 - x - i - 3|1 - i| + 3\lceil x - |x - i|\rceil.$$

The basic diagram to implement logic element $\min(x_1, x_2) \oplus i$ is in Fig. 40.

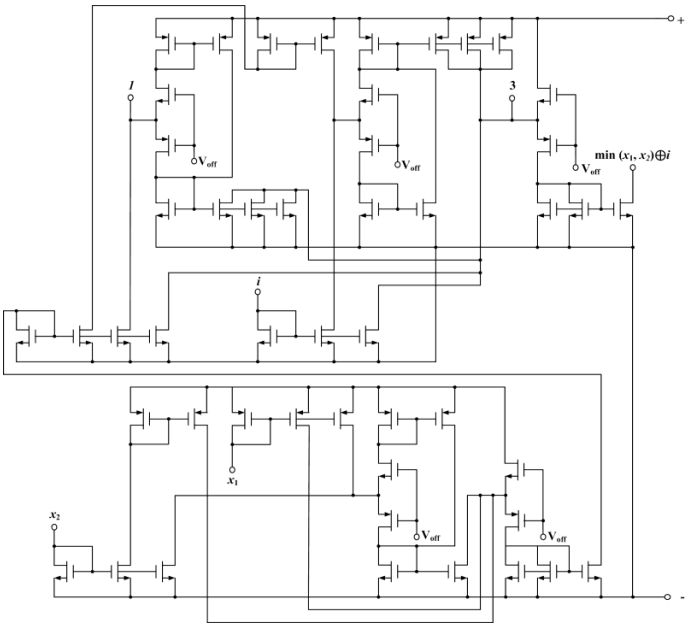


Figure 40: Basic diagram to Implement Difference representation of $\min(x_1, x_2) \oplus i$ operation.

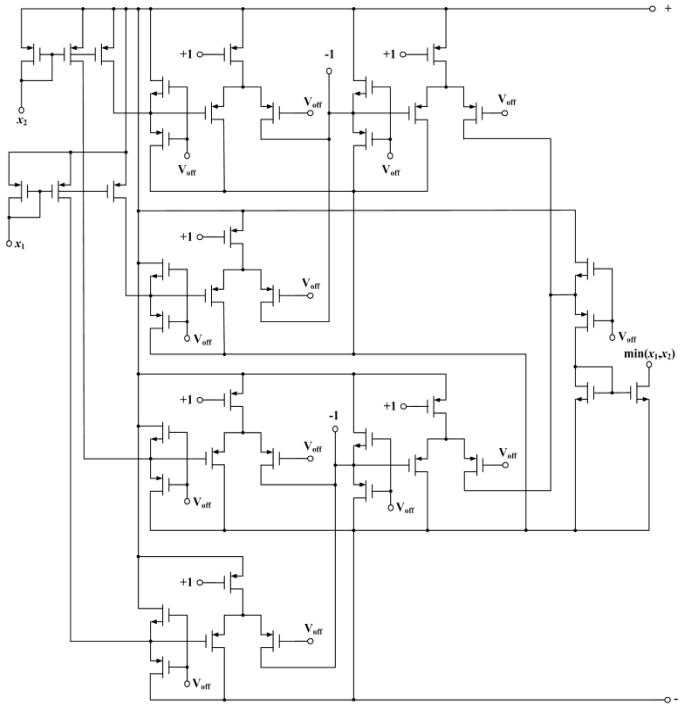


Figure 41. Basic diagram for threshold implementation of $\min(x_1, x_2)$.

$x_1, x_2 \oplus i)$ is an interchange of $\min(x_1, x_2)$ and x⊕i implementations. Similarly, it is possible to implement $\max(x_1, x_2) \oplus i$ and $\max(x_1, x_2 \oplus i)$.

The threshold representation of $\min(x_1, x_2)$ consists of sum of two threshold functions and can be written as following:

$$x_1 > 0) + (x_2 > 0)) > 1] + [((x_1 > 1) + (x_2 > 1)) > 1],$$

Its basic diagram is in Fig. 41.

We can describe x⊕i operation as following

$$x_1 + i + 3[((x_1 > 1) + (i > 1)) > 1] -$$
$$-3[((x_1 > 0) + (i > 1)) > 1] -$$
$$-3[((x_1 > 1) + (i > 0)) > 1].$$

The basic diagram of its implementation is in Fig. 42.

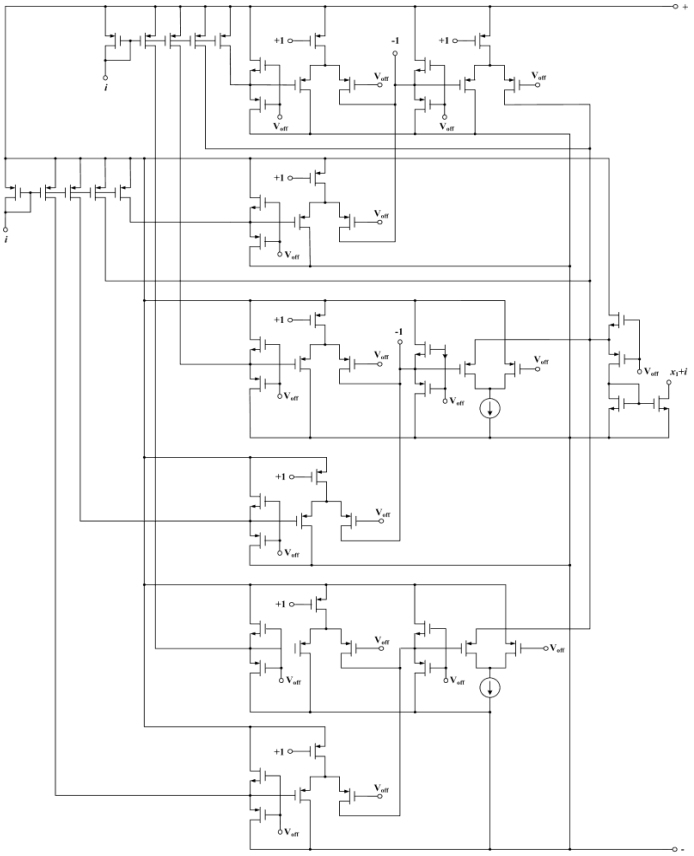


Figure 42: Basic diagram to implement x⊕i .

It is possible to construct logic elements, from which the third-valued flip-flops are constructed, using the above schematic implementations. It is possible to construct flip-flops of different types, following the above structural circuits.

5. Conclusion

The advantages of using the proposed linear circuitry for the construction of flip-flops can be estimated on the basis that:

– all active circuit elements operate in unsaturated mode, therefore, they have higher performance and other frequency characteristics than potential elements;

- circuits contain only active components (transistors), therefore, they are more technological than potential elements; if, for example, to estimate the technological complexity of the elements of LSI contains one conductor, two transistors, three resistors, the technological complexity of the standard two-input potential inverter (11 links, 5 transistors, 5 resistors) is 36, and the linear threshold analog (6 links, 8 transistors) is 22, that is a third less!

- the basis for the construction of the equations of functioning of logic elements is their difference construction, in which the output signal of the element is the difference of the incoming and outcoming components, therefore, linear logic elements have better performance characteristics;

A possible negative indicator of linear circuits may be their energy characteristics: the condition for the operation of the circuits is a constant flow of current in their circuits. However, at high frequencies, these characteristics are likely to differ little from similar characteristics of potential circuits.

Multivalued current logic elements are a new class of elements, which has no analogue in potential logic and there is simply nothing to compare the characteristics of current elements with. Its advantages and disadvantages are all of the above positive and negative qualities of linear circuits, presumably a linear increase in hardware costs for the implementation of schemes with increasing valuedness, as well as the reality of creating a fully functional multi-valued logic elements and digital structures based on them.

Conflict of Interest

The authors declare that there is no conflict of interests regarding publication of this paper.

Acknowledgments

The reported study was funded by RFBR according to the research project No. 18-37-00061.

References

[1] S.V. Yablonsky, “An introduction to mathematical mathematics,” Textbook A manual for universities, 2nd ed, Moscow publ. Nauka, 1986, 384 p. (in Russian).
[2] M.G. Karpovsky, E.S. Moskalev, “Spectral methods of analysis and synthesis of discrete devices,” Moscow, Energy, 1973 – 144 p. (in Russian).
[3] V.D. Efremov, A.A. Kuzmin, V.A. Stepanov, “Calculation of logical functions using the Rademacher transforms,” Automation and Telemechanics, 1984, No. 2. (in Russian).
[4] V.P. Shmerko, “Synthesis of arithmetic forms of Boolean functions by means of the Fourier transform,” Automation and Remote Control. 1989. No. 5. pp. 134-142. (in Russian).
[5] T.I. Post, “Introduction to a genetal theory of elementary propositions,” Math, 1921, 43, No.3, p.163-185. (in Russian).
[6] L. A. Zalmanson, “Fourier transform, Walsh, Haar and their application in management, communications and other fields,” Moscow, Nauka, Fizmatlit publ., 1989, 496 p. (in Russian).
[7] V.L. Artyukhov, V.N. Kondratiev, A.A. Shalyto, “Realization of Boolean functions by arithmetic polynomials,” Automation and Telemechanics. 1988, No. 4, pp. 1-48. (in Russian).
[8] V.D. Malyugin, G.A. Kukharev, V.P. Shmerko, “Transformation of polynomial forms of Boolean functions,” Moscow, Institute of Management Problems, 1986. (in Russian).
[9] V. Malyugin, “Implementation of Boolean functions by arithmetic polynomials,” Automation and Telemechanics, 1982, No. 4, pp. 84–93 (in Russian).
[10] N.N. Prokopenko, N.I. Chernov, V.Ya. Yugai, “Base Concept of Linear Synthesis multi-Valued Digital Structures within Linear Spaces” Proceedings of The IS&IT13 Congress., The Scientific Edition in four volumes, Moscow, PhisMathLit, 2013, vol. 1, pp. 284−289. (in Russian)
[11] E.P. Ugryumov, “Digital circuitry,” BHV Publ, St. Petersburg, 2000, 528 p. (in Russian).
[12] I.N. Bukreev et al, “Microelectronic circuits of digital devices,” Moscow, Radio and communications, 1990, 416 p. (in Russian).
[13] A.F. Gonzalez and P. Mazumder, “Multiple-valued signed digit adder using negative differential resistance devices”, in IEEE Transactions on Computers, vol. 47, no. 9, pp. 947-959, Sept. 1998. DOI: 10.1109/12.713314
[14] A. Morgul and T. Temel, “Current-mode level restoration: circuit for multi-valued logic”, in Electronics Letters, vol. 41, no. 5, pp. 230-231, 3 March 2005. DOI: 10.1049/el:20056998
[15] P. Gonga, “Quaternary logic and applications using Multiple Quantum well based SWSEETs”, IEEE VLSICS, vol. 3, no. 5, pp. 27-42, Oct. 2012. DOI : 10.5121/vlsic.2012.3503
[16] Ch. Babu Dara, T.Haniotakis, S.Tragoudas, “Delay analysis for an N-Input current mode threshold logic gate”, Computer Society Annual Symposium on VLSI 2012 IEEE, DOI: 10.1109/ISVLSI.2012.34
[17] K. Navi, “A Novel Current Mode Full Adder Based on Majority Function”, World Applied Sciences, Journal 4 (5): 676-680, 2008.
[18] N.K. Naware, D.S. Khurge and S.U. Bhandari, “Review of Quaternary Algebra & Its Logic Circuits”, 2015 IEEE ICCUBEA, 2015, pp. 969-973. DOI: 10.1109/ICCUBEA.2015.204
[19] B. Kishor, “Implementation of Ternary / Quaternary Addition using Multivalue Logic Digital Circuit”, International Journal of Computer Application, vol. 118, no. 4, May 2015.
[20] N.N. Prokopenko, N.V. Butyrlagin, N.I. Chernov, V.Ya. Yugay, “Synthesis of binary flip-flops in the apparatus of linear algebra,” News SFU, Technical science, No. 2. 2015, pp. 115-125. (in Russian).
[21] N.N. Prokopenko, N.I. Chernov, V.Ya. Yugay, N.V. Butyrlagin, “Linear synthesis of a k-digit digital element base with current logical signals: the principle of generalization,” Problems of developing promising micro- and nanoelectronic systems, 2016 Proceedings under the general. ed. Academician of the RAS A.L. Stempkovsky, Moscow publ, IPPM RAS, 2016, Part I. P. 70-77. (in Russian).
[22] N.I. Chernov, V.Y. Yugai, N.N. Prokopenko, N.V. Butyrlagin, “Basic concept of linear synthesis of multivalued digital structures in linear spaces,” Proceedings of IEEE East-West Design and Test Symposium, EWDTS 2014, art. no. 7027045. DOI: 10.1109/EWDTS.2014.7027045
[23] N.N. Prokopenko, N.V. Butyrlagin, N.I. Chernov, V.Ya. Yugai, “The linear concept of logical synthesis of digital IP-modules of control and communication systems,” Proceedings of International Siberian Conference on Control and Communications, SIBCON 2015, art. no. 7147182. DOI: 10.1109/SIBCON.2015.7147182.
[24] N.N. Prokopenko, N.V. Butyrlagin, N.I. Chernov, V.Ya. Yugai, “Basic linear elements of k-Valued digital structures,” Proceedings of International Conference on Signals and Electronic Systems, ICSES 2016, pp. 7-12. DOI: 10.1109/ICSES.2016.7847763.
[25] N.N. Prokopenko, N.I. Chernov, V. Yugai, N.V. Butyrlagin, “The element base of the multivalued threshold logic for the automation and control digital devices,” Proceedings of 2017 IEEE International Siberian Conference on Control and Communications, SIBCON 2017, art. no. 7998508. DOI: 10.1109/SIBCON.2017.7998508.
[26] N.N. Prokopenko, N.I. Chernov, V. Yugai, N.V. Butyrlagin, “The multifunctional current logical element for digital computing devices, operating on the principles of linear (not boolean) algebra,” Proceedings of 2016 IEEE East-West Design and Test Symposium, EWDTS 2016, art. no. 7807723. DOI: 10.1109/EWDTS.2016.7807723.
[27] N. Chernov, N. Prokopenko, V. Yugai, N. Butyrlagin "Linear algebra as an alternative approach to the synthesis of digital devices of automation and control systems", Advances in Science, Technology and Engineering Systems Journal, vol. 3, no. 1, pp. 168-190 (2018).

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 441-451 (2020)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Pyrolysis of Beech Wood in A Continuous Drop Tube Reactor in Comparison to A Batch Reactor

Chetna Mohabeer[1,*], Nourelhouda Boukaous[2], Antoinette Maarawi[1], Lokmane Abdelouahed[1], Abdeslam-Hassen Meniai[2], Mustapha Chikhi[2], Bechara Taouk[1]

[1]*Normandie Univ, INSA Rouen Normandie, UNIROUEN, Laboratoire de Sécurité des Procédés Chimiques, LSPC EA-4704, 76000 Rouen, France*

[2]*Faculté de Génie des Procédés, Université de Constantine 3, 25000 Constantine, Algeria*

ARTICLE INFO

Article history:
Received: 07 January, 2020
Accepted: 19 March, 2020
Online: 04 April, 2020

Keywords:
Pyrolysis
Beech wood
Drop tube reactor
Batch spoon-fed reactor

ABSTRACT

The goal of this research work is first to investigate the comparison between a batch and a continuous pyrolysis reactor (at 500, 550 and 600 °C) and the eventual differences and similarities that are obtained by analysing the obtained liquid and gas compositions. All the products formed were examined: liquids by GC-MS/FID and Karl Fischer titration, gases by GC-FID/TCD and solids by thermogravimetric analysis (TGA). It was observed that even though not the same pyrolytic reactor was used, the parameters rendering the maximum yield of bio-oil were the same in both cases: 500 °C under 500 mL/min N_2 (~58 wt. %). Furthermore, it was observed that while changing the temperature affected significantly the product distribution, varying the gas residence time did not cause much change to the product distribution. This was due to the fact that pyrolysis was occurring within the same intermediate regime (τ: ~9 min), and that the limiting factor of the reaction was thus the biomass particle residence time (20.6 s), not that of the gas. Finally, the observations obtained for product distributions and evolution for both the batch and the continuous setups were compared, and discrepancies between results from the two setups were found. This led to the conclusion that comparisons between reactors of different scales need to be undertaken with caution as the results might not be directly comparable.

1. Introduction

This paper is an extension of work originally presented in the 10th International Renewable Energy Congress (IREC) [1]. Depletion of fossil-based fuels has become an important environmental issue. A more sustainable alternative to the use these fuels would be the use of a "green" bio-fuel and a promising route to achieve this goal would be to convert biomass into a product known as "bio-oil" through the pyrolysis process [2]. Lam et al. [3] conducted an exhaustive review that showed that most biomass pyrolysis studies at laboratory scale are performed on batch, or semi-batch, reactors due to the ease which they provide. However, among the different studies led on continuous biomass pyrolysis in literature, another commonly-used technology is the drop tube reactor (DTR), named also the entrained flow reactor or the free fall reactor [3, 4]. This study undertakes a comparison of these two experimental setups.

A number of studies was led on pyrolysis in batch reactors and a pertinent one was that of Bajus [5]. They studied beech wood pyrolysis at temperatures between 350 and 450 °C in a batch reactor and found that the products formed on average were 25.1 wt. % non-condensable gas, 44.3 wt. % liquid and 30.6 wt. % char. The non-condensable gas products comprised mainly of carbon monoxide and carbon dioxide while the liquids contained mostly acetic acid, formic acid, furfural, lactic acid and levoglucosan. Another study led by Demirbas [6] examined the effect of varying reaction temperature (350-525 °C) on the pyrolytic products of hazelnut shell, olive husk, beech and spruce wood samples. The experiments were performed in a fixed-bed tubular reactor. They observed that a significant oil fraction contained phenols and that the water content oscillated around 15-30 wt. % of the total oil mass. They also found that the bio-oil recovered at 500 °C included elevated concentrations of acetic acid, 1-hydroxy-2-butanone, 1-hydroxy-2-propanone, methanol, 2,6-dimethoxyphenol, 4-methyl-2,6-dimethoxyphenol and 2-cyclopenten-1-one.

*Chetna Mohabeer, 685 Avenue de l'Université, 76800 Saint-Etienne-Du-Rouvray, France, +33232956667, chikirsha.mohabeer@insa-rouen.fr

Concerning studies led on drop-tube reactors, Guizani et al. [7] conducted experiments using a free-fall reactor by varying operational conditions for biomass pyrolysis: temperature (450-600 °C), biomass particle size (370-640 µm) and gas residence time (12.6-20.6 s). They firstly observed that more elevated temperatures, along with more elevated particle sizes, caused a decrease in bio-oil yield. Among the various operating conditions, it was seen that temperature of the reactor and the particle size impacted greatly the product distribution without influencing notably properties of the bio-oil, namely, acidity. Sun et al. [8] conducted another interesting study on this subject where flash pyrolysis of rice husk and sawdust was performed in a free fall reactor within a temperature range of 700 to 1000 °C. Results from their study have demonstrated that temperature affected the reaction significantly; when temperature increased, gas yield increased as well while liquid and solid yields lessened. The authors also noted that at less elevated temperatures, the improvement in gas production could be owing to a boost in the formation of CO and CH_4 whereas, at higher temperatures, it concerned increased CO_2 and H_2 production.

Based on the values obtained for residence time for this study, it can be inferred that the pyrolysis regime used was an intermediate one. Now, as stated in literature [9–13], results obtained from fast and intermediate regimes are similar to each other. The latter regime was also the one obtained from the pyrolysis experiments in the batch, spoon-fed reactor (τ: ~9 min based on heated length of reactor). As the same pyrolysis regime reigned in both setups, the results are comparable in terms of operating conditions (like biomass used, reaction temperature and gas residence gas).

This research work presents liquid and gas results obtained for pyrolysis reactions conducted in a batch, spoon-fed reactor (SFR) and a continuous DTR, respectively. This research is a follow-up to our previous studies [13–15] and many of the comparisons makes use of the results derived from these studies. The aim is firstly to establish a comparison between the two reactors and then, study the subsequent similarities and differences obtained from the products of the two setups. To further corroborate these observations, the results were also compared to those obtained from a micro pyrolyser [16]. To the best of our knowledge, such a comparison has rarely been undertaken before, despite the importance of knowing and understanding the variation in results obtained when upscaling a process (changing reactor technology and scale of the process). Ultimately, this study can help attest whether comparing results obtained from different reactor technologies is reliable or not, and why.

2. Experimental Section

2.1. Materials used

The biomass used for this study, beech wood having an average particle size of 400 µm, was supplied by ETS Lignex Company. The elemental and proximate analyses for the biomass can be found in the supplementary materials (Tables A.1 and A.2). The C, H and N contents (elemental analysis) of the samples were measured by the use of a CHN elemental analyser (Flash 2000, Thermofisher Scientific). The proximate analysis was obtained via thermogravimetric analysis (TGA), using a STD Q600 equipment, according to the protocol described by Garcia et al. [17].

2.2. Experimental and analytical setup

2.2.1 Experimental setup

Previous studies conducted by our research group detail the reactor layout, experimental protocol and analytical procedures undertaken for experimental runs in the spoon-fed reactor [13–15]. Now, the experimental runs in continuous mode were performed in a drop tube reactor, as is illustrated by Figure 1. Biomass was fed through the top of the DTR by a screw conveyor (GLD 77.0, Tremie 5L from Gericke). A 2 m heated length of the pyrolysis reactor was used. A char collector to gather solid residues of biomass pyrolysis was placed at the bottom of the DTR. This collector was kept at 300 °C through the use a heating shell to ensure that the condensable pyrolysis vapours entering it stayed in vapour form. However, sometimes, some fine solid particles got swept away by the gas/vapour flow and escaped the char collector. Two cyclones (efficiency: 85 % each) were positioned at the vapour exit of the DTR so as to capture these particles away from the gas flow. It should be noted that the whole system was kept at 300 °C in order to prevent condensation of vapours.

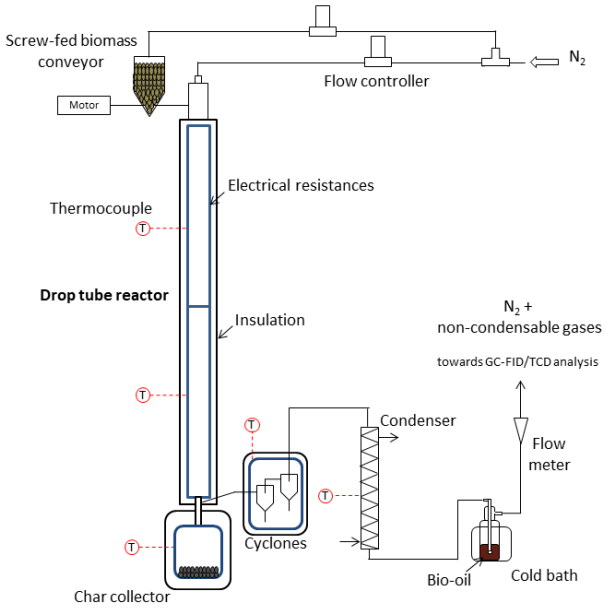


Figure 1: DTR layout

The pyrolysis experiments were performed by altering the temperature of the DTR (500, 550 and 600 °C) and the nitrogen flow rate (0.5, 1 and 2 L/min) in order to vary the gas residence time (τ). The list of experiments has been presented in Table 1. It should be noted that the gas residence time has been calculated based only on the nitrogen flow rate. The experimental protocol followed was: first, the reactor was set to the temperature desired and the nitrogen flow rate was controlled at the exit to make sure that there were no leaks. Then, the system was allowed to stabilise for about two hours. Afterwards, a biomass feed rate of 2 g/min was introduced in the reactor and the reaction was left to occur for 60 minutes. After the experiment, the heat was cut and the experimental system left to cool. The collection of char was done at temperatures below 150 °C so as to prevent it from burning at the contact of air. The collection of bio-oil was done by using a solvent: acetone with 99.98% purity, with, as internal standard, a known added amount of nonane.

Table 1: List of pyrolysis experiments and corresponding gas residence times

Nitrogen flow rate (L/min)*	Gas residence time, τ (min)	DTR temperature (°C)
0.5	9.27	500
		550
		600
1	4.63	500
		550
		600
2	2.32	500
		550
		600

*Measured at room temperature and pressure

2.2.2 Analytical setup for pyrolytic products

Gas Chromatography (GC) Analyses

The liquid and gaseous products formed were then recovered and analysed using the protocol described by our previous work [13]. The bio-oil collected was analysed using a gas chromatograph-mass spectrometer instrument GC-MS (Varian 3900-Saturn 2100T) with a VF-1701ms (Agilent) (60 m × 0.25 mm × 0.25 µm film thickness) column. The same temperature programme as Charon et al. [18] was used. The identification of the liquid components was done using Varian WS (WorskStation) and NIST 2002 software. A GC-FID (flame ionisation detector) Scion 456-GC Bruker instrument was used to quantify the compounds once the identity of the peaks was obtained and confirmed. The same column as for the GC-MS was used. The temperature programme used was the same as for the GC-MS, only, the final temperature used was 240 °C instead of 280 °C. The obtained compounds were finally classified into chemical "families", each having the same main functional group.

The non-condensable gases (NCG) were mostly composed of CO, CO_2, CH_4 and H_2. C_2H_2, C_2H_4 and C_2H_6 (represented as 'C_2') and C_3H_4, C_3H_6 and C_3H_8 (represented as 'C_3') were also present in less significant proportions. The gases were analysed with a gas chromatograph instrument (Clarus 580, Perkin Elmer), equipped with an FID and a thermal conductivity detector (TCD). A Shincarbon St 100 120 column, a methaniser and a hydrogen generator completed the build of the equipment.

2.3 Karl Fischer (KF) titration

A Metrohm 870 KF TitrinoPlus apparatus was used to perform volumetric KF titrations. An aqualine sodium tartrate solution along with Hydranal Composite 5 were used as titrating agents and Hydranal Methanol Rapid was used as working medium. The pyrolysis oil samples were weighed on an analytical precision balance. After weighing, the sample was placed into the titration vessel of the KF apparatus. The titration started when the sample was completely dissolved. Each sample was measured in triplicate, which allowed the calculation of a 99 % confidence interval.

2.4 Thermogravimetric analysis (TGA)

Weight loss as a function of temperature was recorded using a thermobalance (TA Instrument Explorer Q600) from 20 °C to 800 °C. In a typical experiment, 5-10 mg of sample were placed in an alumina crucible and heated under a N_2 flow (50 mL/min). A heating rate of 15 °C/min was used.

3 Results and discussion

3.1 Results from drop tube reactor (DTR)

The major chemical compounds identified and the different families they were grouped into have been summarised in Table 2.

Table 2: Major chemical compounds identified and their different chemical families

Chemical family	Major molecules identified
Alcohols	1,2-Propanediol, 3-methoxy-, 2-Furanmethanol, 6-Hepten-1-ol, 5-methyl-, 2,4-Hexadien-1-ol, Resorcinol, 3-Heptanol, 3,6-dimethyl-, Thymol, 1,2-Benzenediol
Aldehydes	2-Heptenal, (E)-, 2-Pentenal, (E)-, 3-Furaldehyde, 2-Furancarboxaldehyde, 5-methyl-, Benzaldehyde, 2-hydroxy-
Alkanes	Hexane, 3,3,4,4-tetramethyl-, Nonane, Decane
Aromatics	Benzene, o-Xylene, Benzene, 1,3-dimethyl-, 1H-Indene, 3-methyl-, Naphthalene, 1,2-dihydro-, Azulene, Naphthalene, 2,6-dimethyl-
Amides	N-Methoxy-N-methylacetamide, Benzenacetic amide
Carboxylic acids	Acetic acid, Acetohydroxamic acid, Propanoic acid, Benzoylformic acid
Esters	Acetic acid, 1-methylethyl ester, 1,1-Ethanediol, diacetate, Phenol, 4-(2-propenyl)-, acetate
Furans	Furan, 2,3,5-trimethyl-, Furan, 2-ethyl-, Vinylfuran, Benzofuran, 2-methyl-
Guaiacols	Phenol, 2-methoxy-, Phenol, 2,6-dimethoxy-, Phenol, 4-ethyl-2-methoxy-
Ketones	1-Hexen-3-one, 2-Hexanone, 3,4-dimethyl-, 2-Cyclohexen-1-one, 2-Cyclopenten-1-one, 1H-Inden-1-one, 2,3-dihydro-
Phenols	Phenol, Phenol, 2-methyl-, Phenol, 2,4,6-trimethyl-, Phenol, 4-ethyl-, 1-Naphthalenol
Carbohydrates	Levoglucosan, 3,4-O-Isopropylidene-d-galactose

3.2 Defining optimal conditions for pyrolysis reaction in the DTR

The mass balance values obtained for the various experimental runs are presented in Figure 2. It is worth mentioning that mass balances of above 85 % were found experimentally for each run. This discrepancy can be explained by the fact that there still remained a fraction of the liquid produced on the walls of the DTR installation that was lost even if every attempt was made to collect all of it. Therefore, the mass balances were completed to 100 % with respect to the liquid fraction.

It can firstly be observed from the previous figure that the lowest temperature (500 °C) tested gave rise to the largest amount

of char produced (between 13 and 18 wt. %) and the smallest amount of non-condensable gases formed (between 28 to 30 wt. %). In contrast to the latter observations, 600 °C, the highest temperature utilised corresponded to the lowest amount of char (around 9 to 10 wt. %) and the highest amount of non-condensable gases (49-55 wt. %). Now, so as to detail these findings, the degree of conversion of the biomass at different temperatures, that is, the fraction of biomass that was converted to its pyrolytic products at each pyrolytic temperature, was analysed through TGA under a nitrogen flow rate of 0.5 L/min and the results obtained have been illustrated in Figure 3. The figure also depicts the evolution of the percentage of each pyrolytic product at different reactor temperatures.

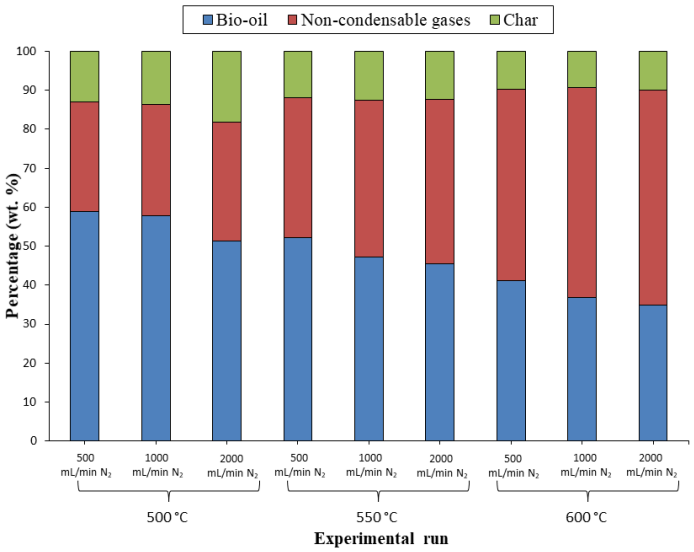


Figure 2: Mass balances of pyrolysis experiments

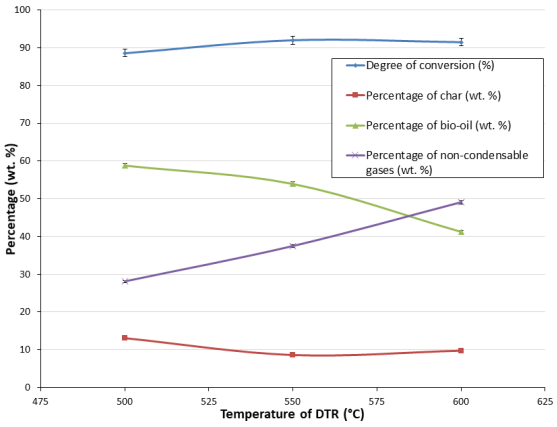


Figure 3: Degree of biomass conversion (TGA) at different DTR temperatures at τ = 9.27 min

It can be seen from the previous figure that the tendency of the degree of conversion and that of the char percentage were opposing ones, as were the evolution of bio-oil and gas percentages. It can therefore be claimed that the gas production bears a direct correlation to the reactor temperature, and that the liquid formation a negative one, that is, at higher temperatures, an improvement in gas formation and a diminishment in liquid production will be witnessed. This observation can be explained by the fact that cracking reactions are favoured at higher temperatures [19], and so, at those temperatures, larger liquid molecules decompose into smaller, gaseous ones. The same observations were obtained from the SFR, as mentioned in previous works [13, 14].

It was also observed that the experiment performed at 500 °C under 0.5 L/min N_2 produced the most elevated bio-oil output (58.8 wt. %). These conditions corresponded to the same ones obtained in the SFR [14].

3.3 Influence of gas residence time on pyrolysis liquid and gas products

3.3.1 Liquid product composition

As previously mentioned, in order to elicit a variation in the gas residence time in the DTR, carrier gas (N_2) flow rate was changed from 0.5 to 2 L/min. Table 1 shows the resulting residence times. Then, the liquid product compositions for the pyrolysis of beech wood obtained at different reactor temperatures were analysed. Figure 4 illustrates the composition obtained at 500 °C in the DTR. The ones obtained for other reactor temperatures have been tabulated in Table 3.

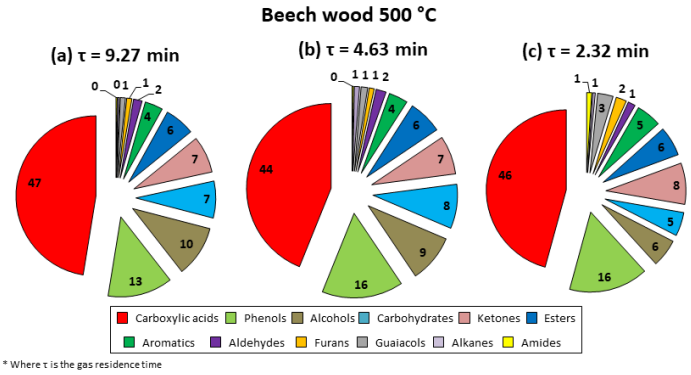


Figure 4: Liquid product compositions obtained for pyrolysis of beech wood at 500 °C in DTR

As it can be seen in Figure 4, although gas residence times have more than halved (9.27 min to 4.63 min to 2.32 min), product compositions were not much influenced. Further calculations were conducted to try to understand the happenings throughout the reaction. These calculations have been compiled in the supplementary materials. From these calculations, it was found that the biomass particle residence time in the reactor was 2.20 s, which was much lower than the gas residence times used (see Table 2). This observation led to the understanding that the parameter that controlled the formation of products and the dynamics of the reaction, that is, the limiting parameter, was in fact the biomass particle residence time, not the gas residence time. Thus, even if the carrier gas flow were increased, hence diminishing gas residence time, the liquid product distribution would undergo no appreciable change as long as the pyrolysis reaction took place in a regime wherein the biomass particle residence time was less than the gas residence time. The same observation was also made by Guizani et al. [7], Jahirul et al. [20] and Ellens [21]. Now, it could be stipulated that raising the N_2 flow rate to higher values would help to overcome this obstacle.

However, for this specific setup, technical restrictions did not allow the flow rate used to be any higher than 2 L/min.

3.4 Gas product composition

NCG were also examined while the N_2 flow rate was altered. Figure 5 illustrates the different NCG compositions obtained at 0.5, 1 and 2 L/min N_2 at 500 °C, while the compositions at other temperatures have been listed in the supplementary materials (Table A.3).

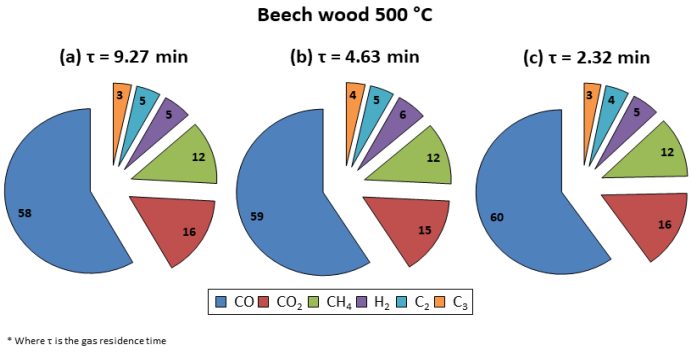


Figure 5: Gas product compositions obtained for pyrolysis of beech wood at 500 °C in DTR

As can be observed from Figure 5, the pattern for the gas composition resembled that for the liquid composition: the gas composition was not majorly influenced by elevating the carrier gas flow rate. Following this, the same inference can be made: even if the gas residence time changes, the gas product distribution is not majorly affected if the pyrolysis regime is not altered.

3.5 Comparison of results from the SFR and the DTR

3.5.1 Liquid product composition

To enhance the visual representation of the various chemical families, pie charts were constructed for the different temperatures and nitrogen flow rates used. An example is Figure 6 (composition of beech wood bio-oil obtained at 500 °C under 0.5 L/min N_2 in DTR compared to that obtained in SFR under same conditions). A similar analysis was performed for each bio-oil produced (at 500, 550 and 600°C under N_2 flow rates of 0.5, 1 and 2 L/min) and the results have been presented in Table 3.

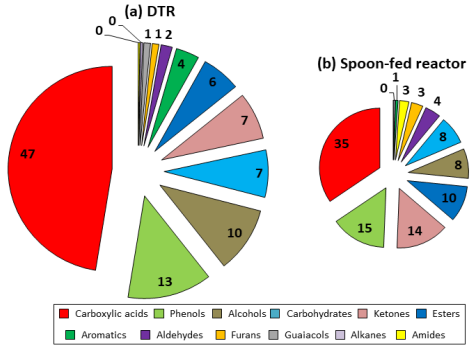


Figure 6: Composition of beech wood bio-oils obtained at 500 °C and τ = 9.27 min in DTR and spoon-fed reactor (τ = 9 min) (mol. %)

It can be observed from Figure 6 that carboxylic acids, phenols and alcohols were the most important chemical families present in the oils. These findings concord to a certain extent with the results found when the SFR was used; the only dissimilarity was that the oils obtained from the SFR contained a higher quantity of ketones and esters. This difference may be explained by the variation in the collected oil components: concerning the DTR, using a more important amount of biomass led to a more significant production of bio-oil, which demanded the use of a smaller amount of solvent so as to be collected. This enabled the detection of compounds which had been present only in negligible fractions in the oil obtained from the SFR, which had been more dilute (338 detected compounds vs. 255 in spoon-fed reactor). The observed variation may also be due to the change in reaction technology used. However, these results vary slightly from those that Guizani et al. [7] obtained. They investigated the fast pyrolysis of biomass in an entrained flow reactor by changing reactor temperature, size of particles and vapour residence time. The major oil components observed by them were aldehydes, ketones and carboxylic acids. This difference might stem from the fact that even though beech wood was used as raw material for their study, they used a gas residence time of 20.6 s for a mean biomass particle size of 640 µm, which both differ from the conditions used in this study. Nevertheless, their maximum oil yield, 62.4 wt. %, was found to be formed at 500 °C, which this time, coincided with the findings presented in this study.

Now, it is interesting to study increasing reactor temperature influences the major families present in oil (the trend has been illustrated in Figure 7). The influence on all the other chemical families has been depicted in the supplementary materials, Figures A.1 (a)-(l).

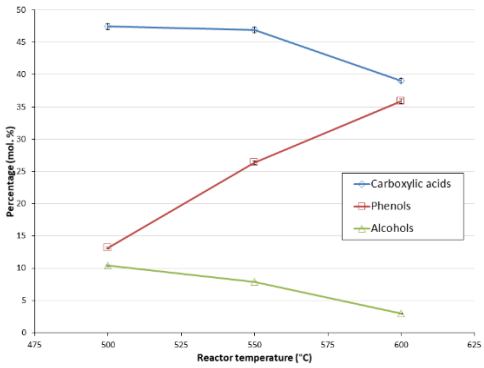


Figure 7: Evolution of major bio-oil components with increasing DTR temperature (τ: 9.27 min)

Figure 7 shows clearly that both carboxylic acid and alcohol percentages decreased when temperature increased, while the opposite trend was portrayed by phenols. These findings do not completely match the results from the spoon-fed reactor, where an increasing trend with temperature was noted for the carboxylic acids.

The diminishment in the percentage of carboxylic acids corresponded with a sharp increase in the percentage of phenol. As mentioned in our previous work [13], carboxylic acids and phenols tend to be negatively correlated; this means that, when one experiences an increase, the other undergoes a reduction. This observation has been well-depicted in Figure 7. Now, a certain difference from results obtained from the SFR has been observed.

Table 3: Chemical family percentages (mol. %) in bio-oil samples obtained from DTR

Experimental conditions	Chemical families present (mol. %)											
	Carboxylic acids	Alkanes	Aromatics	Alcohols	Aldehydes	Amides	Ketones	Esters	Furans	Guaiacols	Phenols	Carbohydrates
500 °C 0.5 L/min N_2	47.44	0.36	3.64	10.38	1.74	0.25	7.45	6.10	1.04	1.02	13.12	7.45
500 °C 1 L/min N_2	43.88	1.00	3.65	9.18	1.94	0.26	7.38	6.48	0.97	1.26	15.55	8.46
500 °C 2 L/min N_2	45.75	0.53	5.09	5.67	1.49	1.01	8.35	5.99	2.13	3.15	16.03	4.81
550 °C 0.5 L/min N_2	46.91	0.55	2.95	4.04	1.75	0.90	5.94	5.00	1.02	1.05	26.33	3.57
550 °C 1 L/min N_2	46.34	0.29	2.60	7.92	1.40	0.72	7.24	3.28	1.00	1.04	24.88	3.29
550 °C 2 L/min N_2	47.90	1.59	3.63	5.76	1.66	0.15	6.56	3.88	0.98	1.07	23.74	3.08
600 °C 0.5 L/min N_2	39.04	0.33	4.90	2.97	1.76	2.08	4.75	5.21	0.71	0.98	35.90	1.38
600 °C 1 L/min N_2	38.35	0.21	2.80	3.60	1.31	1.98	4.78	8.59	1.26	1.21	34.71	1.21
600 °C 2 L/min N_2	39.75	0.34	3.50	3.37	1.69	2.53	5.66	4.37	0.82	1.18	35.09	1.69

This variance can be explained firstly by the fact that a higher number of molecules was detected for the DTR and secondly, by the operational conditions of the two reactors used, which varied significantly. This finding illustrates that changing a reaction scale from lab to pilot may ensue in differences in the results obtained, and the same may be observed when reaction technology is changed from batch or semi-continuous to continuous. The latter statement is further corroborated by Table 4, where results obtained from different pyrolysis technologies used have been presented to better grasp the similarities and differences.

Table 4: Comparison of different pyrolysis technologies

		μ-pyrolyser	SFR	DTR
	Amount of raw material	~mg	~g	~100 g
Liquid fraction	No. of molecules identified	92	255	338
	Major compounds	40-46 % acetic acid No levogluco-san	35 % acetic acid 8 % levogluco-san	47 % acetic acid 7 % levogluco-san
	Trend of acetic acid with reactor temperature	↘	↗	↘
Gas fraction		No gas analysis performed as only traces were detected	Detection of: 45 % CO 40 % CO_2 11 % CH_4 3 % H_2	Detection of: 58 % CO 16 % CO_2 12 % CH_4 5 % H_2
References		[16]	[13, 14]	This study

3.6 Gas product composition

NCG went through the same analysis that liquid products did. Figure 8 represents the gas product distributions obtained for samples at 500 °C under 0.5 L/min N_2 in the DTR and the SFR. Results compiled from diagrams showing NCG compositions when reaction conditions are varied have been listed in Table A.3, found in the supplementary materials.

Firstly, it was observed that CO was the majorly present species (58 vol. %). This result concords with that of Guizani et al. [7] (58.4 vol. %), and the variance from what was found with the SFR was not very large (45 vol. % for CO, the major gas component). However, it was witnessed that the gaseous product distribution changed slightly with DTR use: a more important fraction of each NCG component was detected. This may be owing to the fact that, as it was with the liquids, a higher amount of biomass was supplied to the DTR than to the SFR and this led to the production of more concentrated NCG, enabling easier detection of smaller gas molecules that were henceforth less dilute in the carrier gas. Furthermore, Figure 9 was obtained by analysing the trend of each NCG component at 0.5 L/min N_2 with increasing reactor temperature. The graphs obtained with the use of other flow rates have been presented in the supplementary materials, Figures A.2 and A.3.

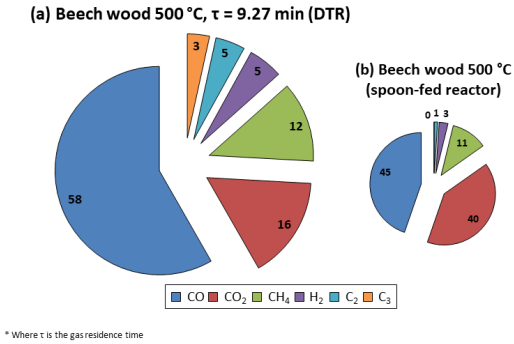


Figure 8: Composition of beech wood NCG obtained at 500 °C in DTR and spoon-fed reactor (vol. %)

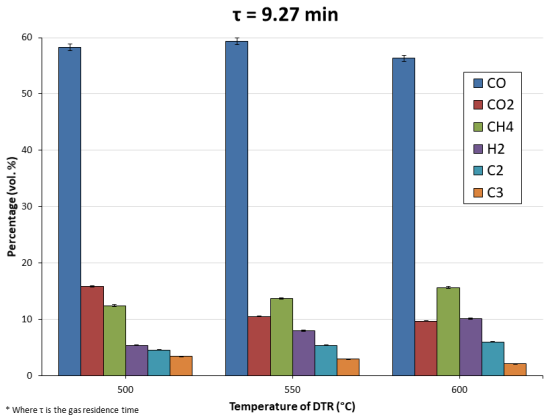


Figure 9: Evolution of NCG components (vol. %) with DTR temperature at τ = 9.27 min

The NCG composition evolved quite feebly at different DTR temperatures: a spike of CO was noted at 550 °C while increasing pyrolysis temperature caused a CO_2 percentage decrease. All other gaseous compounds, excluding C_3, went through a raise in percentage with increasing reactor temperature. The general increase in CO and decrease in CO_2 trend observed with the SFR when reaction temperature was increased, somewhat shifted when the DTR was used. In this case, it was seen that whilst a decrease in the percentage of CO_2 was indeed witnessed, CO percentage remained almost stable (as CO was the major component). Still, if quantitative values are taken into consideration, a raise in the volume of CO can be observed, substantiating the theory that mounting the temperature of the reactor indeed privileged cracking reactions and thereby, formation of gas molecules, as per what was mentioned previously.

3.7 Water content of liquid products

KF titration method was utilised to estimate water content for each experiment. The detailed results can be found in Figure 10. It can be observed that a reduction in the water content was caused by a raise in the temperature of DTR (maximum water content: 48 wt. % at 600 °C under 1 L/min N_2). This coincided with a

reduction in some chemical families such as alcohols, carbohydrates and carboxylic acids as well. Similar findings were made by Lehto et al. [22]. It can be therefore be claimed that elevating the temperature of the reactor promoted the dehydration of the previously-mentioned chemical families.

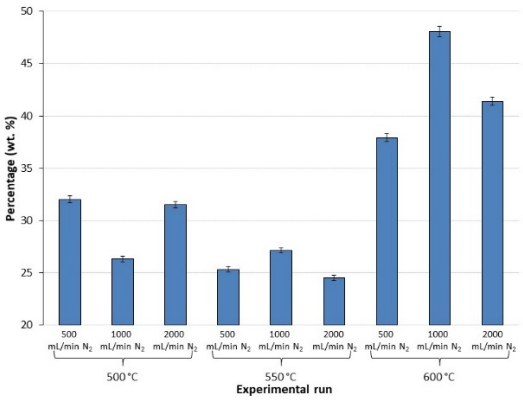


Figure 10: Water content (wt. %) of bio-oils obtained from each experimental run

3.8 Oxygen content of liquid products

The elevated oxygen fraction of pyrolytic oils, which renders them highly functionalised [23], remains a crucial challenge to their mainstream use. This high functionality is what causes the bio-oils to be relatively highly viscous, acidic and unstable and to possess low calorific values. Therefore, acquiring an indication of the oxygen fraction formed directly from the compounds found in the bio-oil (oxygen fraction derived from humidity not included) is key to be able to understand the behaviour of the oil and therefore, find a way to improve its properties. The theoretical calculation of the oxygen content was conducted from the weight of the oxygen atoms present in the compound. Figure 11 illustrates the evolution of the bio-oils' oxygen content (mol. %) obtained at DTR different temperatures under various N_2 flow rates.

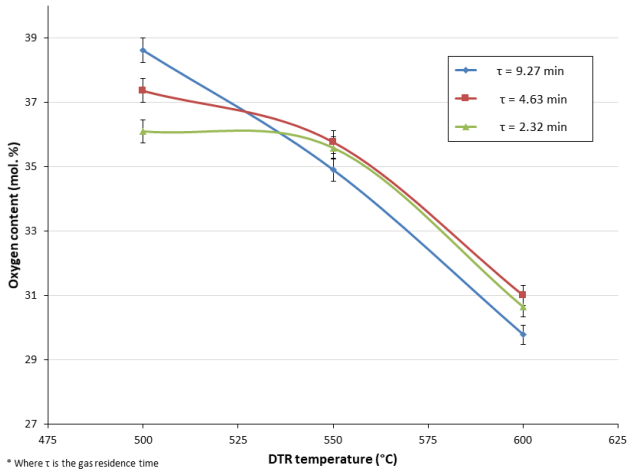


Figure 11: Oxygen content (mol. %) of bio-oils obtained from all experimental runs

It can be observed that increasing the temperature of the DTR caused the oxygen content to lessen. This finding corroborates the fact that cracking reactions and the elimination of certain oxygen-containing chemical groups are favoured by more elevated reaction temperatures. Thus, the formation of smaller, gaseous molecules was privileged. Hence, at more elevated temperatures, a less significant amount of liquid was collected and some of the oxygen atoms escaped in the gas fraction. Nonetheless, the global bio-oil oxygen content remained quite high (29-40 mol. %) in contrast to that of the bio-oils obtained from the spoon-fed reactor (average of 33 mol. %) and also, that of fossil-derived fuels (~1 %) [24].

4 Conclusion

The results pertaining to liquid and gas products obtained for a continuous pyrolysis reaction in a DTR were presented in this study. The goal of this research work was first to investigate a comparison a batch and a continuous pyrolysis reactor and the subsequent existing differences and similarities by examining the obtained liquid and gas compositions. The temperature of the DTR and the gas residence time were primarily changed to study the effect on the pyrolytic products. Even though not the same pyrolytic reactor was used, it was observed that the parameters rendering the maximum yield of bio-oil were the same in both cases: 500 °C under 500 mL/min N_2 (~58 wt. %). It was also observed that while changing the temperature affected significantly the product distribution (it was seen that the use of higher temperatures gave rise to higher amounts of gas products and a more acidic oil), changing the gas residence time did not cause much change to the product distribution. This was due to the fact that pyrolysis was occurring within the same intermediate regime (τ: ~9 min), and that the reaction limiting factor was the residence time of the biomass particle (20.6 s), not that of the gas. Finally, many inconsistent findings were observed between the spoon-fed and the continuous setups. This means that when comparisons are undertaken for results obtained from setups of different scales, it should be with caution as the results might not be directly comparable. This fact highlights the need for deeper research so as to understand the differences in the results obtained from differently-scaled reactors.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This project has been funded with the support from the European Union with the European Regional Development Fund (ERDF) and the Regional Council of Normandie.

References

[1] N. Boukaous, L. Abdelouahed, M. Chikhi, A.-H. Meniai, and B. Taouk, 'Combustion behaviour of several Mediterranean biomass sources and their blends biomass/char', in *2019 10th International Renewable Energy Congress (IREC)*, 2019, pp. 1–6, doi: 10.1109/IREC.2019.8754608.

[2] G. Gentile, A. Cuoci, A. Frassoldati, T. Faravelli, and E. Ranzi, 'A comprehensive cfd model for the biomass pyrolysis', *Chemical Engineering Transactions*, pp. 445–450, 2015, doi: 10.3303/CET1543075.

[3] S. S. Lam, R. K. Liew, A. Jusoh, C. T. Chong, F. N. Ani, and H. A. Chase, 'Progress in waste oil to sustainable energy, with emphasis on pyrolysis

techniques', *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 741–753, Jan. 2016, doi: 10.1016/j.rser.2015.09.005.
[4] L. Lu, Y. Jin, M. R. Nakamura, M. J. Castaldi, and K. Yoshikawa, 'An Investigation on Cocombustion Behaviors of Hydrothermally Treated Municipal Solid Waste with Coal Using a Drop-Tube Reactor', *Journal of Combustion*, 2012. [Online]. Available: https://www.hindawi.com/journals/jc/2012/282851/. [Accessed: 20-Sep-2018].
[5] M. Bajus, 'Pyrolysis of woody material', *Petroleum & Coal*, vol. 52, no. 3, pp. 207–214, 2010.
[6] A. Demirbas, 'The influence of temperature on the yields of compounds existing in bio-oils obtained from biomass samples via pyrolysis', *Fuel Processing Technology*, vol. 88, no. 6, pp. 591–597, Jun. 2007, doi: 10.1016/j.fuproc.2007.01.010.
[7] C. Guizani, S. Valin, J. Billaud, M. Peyrot, and S. Salvador, 'Biomass fast pyrolysis in a drop tube reactor for bio oil production: Experiments and modeling', *Fuel*, vol. 207, pp. 71–84, Nov. 2017, doi: 10.1016/j.fuel.2017.06.068.
[8] S. Sun, H. Tian, Y. Zhao, R. Sun, and H. Zhou, 'Experimental and numerical study of biomass flash pyrolysis in an entrained flow reactor', *Bioresource Technology*, vol. 101, no. 10, pp. 3678–3684, May 2010, doi: 10.1016/j.biortech.2009.12.092.
[9] A. Hornung, 'Intermediate Pyrolysis as an Alternative to Fast Pyrolysis', *BioEnergy IV: Innovations in Biomass Conversion for Heat, Power, Fuels and Chemicals*, Jun. 2013.
[10] A. S. N. Mahmood, J. G. Brammer, A. Hornung, A. Steele, and S. Poulston, 'The intermediate pyrolysis and catalytic steam reforming of Brewers spent grain', *Journal of Analytical and Applied Pyrolysis*, vol. 103, pp. 328–342, Sep. 2013, doi: 10.1016/j.jaap.2012.09.009.
[11] I. D. V. Torri *et al.*, 'Bio-oil production of softwood and hardwood forest industry residues through fast and intermediate pyrolysis and its chromatographic characterization', *Bioresour. Technol.*, vol. 200, pp. 680–690, Jan. 2016, doi: 10.1016/j.biortech.2015.10.086.
[12] K. Kebelmann, A. Hornung, U. Karsten, and G. Griffiths, 'Intermediate pyrolysis and product identification by TGA and Py-GC/MS of green microalgae and their extracted protein and lipid components', *Biomass and Bioenergy*, vol. 49, pp. 38–48, Feb. 2013, doi: 10.1016/j.biombioe.2012.12.006.
[13] C. Mohabeer, L. Abdelouahed, S. Marcotte, and B. Taouk, 'Comparative analysis of pyrolytic liquid products of beech wood, flax shives and woody biomass components', *Journal of Analytical and Applied Pyrolysis*, vol. 127, pp. 269–277, Sep. 2017, doi: 10.1016/j.jaap.2017.07.025.
[14] C. Mohabeer *et al.*, 'Production of liquid bio-fuel from catalytic de-oxygenation: Pyrolysis of beech wood and flax shives', *Journal of Fuel Chemistry and Technology*, vol. 47, no. 02, pp. 153–166, Feb. 2019.
[15] C. Mohabeer, L. Reyes, L. Abdelouahed, S. Marcotte, and B. Taouk, 'Investigating catalytic de-oxygenation of cellulose, xylan and lignin bio-oils using HZSM-5 and Fe-HZSM-5', *Journal of Analytical and Applied Pyrolysis*, vol. 137, pp. 118–127, Jan. 2019, doi: 10.1016/j.jaap.2018.11.016.
[16] J. D. Murillo, E. A. Ware, and J. J. Biernacki, 'Characterization of milling effects on the physical and chemical nature of herbaceous biomass with comparison of fast pyrolysis product distributions using Py-GC/MS', *Journal of Analytical and Applied Pyrolysis*, vol. 108, pp. 234–247, Jul. 2014, doi: 10.1016/j.jaap.2014.04.010.
[17] L. Garcia, M. L. Salvador, J. Arauzo, and R. Bilbao, 'Influence of catalyst weight/biomass flow rate ratio on gas production in the catalytic pyrolysis of pine sawdust at low temperatures', *Ind. Eng. Chem. Res.*, vol. 37, pp. 3812–3819, 1998.
[18] N. Charon *et al.*, 'Multi-technique characterization of fast pyrolysis oils', *Journal of Analytical and Applied Pyrolysis*, vol. 116, pp. 18–26, Nov. 2015, doi: 10.1016/j.jaap.2015.10.012.
[19] Y. Fan, Y. Cai, X. Li, N. Yu, and H. Yin, 'Catalytic upgrading of pyrolytic vapors from the vacuum pyrolysis of rape straw over nanocrystalline HZSM-5 zeolite in a two-stage fixed-bed reactor', *Journal of Analytical and Applied Pyrolysis*, vol. 108, pp. 185–195, Jul. 2014, doi: 10.1016/j.jaap.2014.05.001.
[20] M. I. Jahirul, M. G. Rasul, A. A. Chowdhury, and N. Ashwath, 'Biofuels Production through Biomass Pyrolysis —A Technological Review', *Energies*, vol. 5, no. 12, pp. 4952–5001, Nov. 2012, doi: 10.3390/en5124952.
[21] C. Ellens, 'Design, optimization and evaluation of a free-fall biomass fast pyrolysis reactor and its products', *Graduate Theses and Dissertations*, Jan. 2009.
[22] J. Lehto, A. Oasmaa, and Y. Solantausta, 'Fuel oil quality and combustion of fast pyrolysis bio-oils'. 2013.
[23] J. D. Adjaye and N. N. Bakhshi, 'Production of hydrocarbons by catalytic upgrading of a fast pyrolysis bio-oil. Part I: Conversion over various catalysts', *Fuel Processing Technology*, vol. 45, no. 3, pp. 161–183, Dec. 1995, doi: 10.1016/0378-3820(95)00034-5.
[24] S. Xiu and A. Shahbazi, 'Bio-oil production and upgrading research: A review', *Renewable and Sustainable Energy Reviews*, vol. 16, no. 7, pp. 4406–4414, Sep. 2012, doi: 10.1016/j.rser.2012.04.028.
[25] Engineering ToolBox, 'The Engineering ToolBox', 2001. [Online]. Available: https://www.engineeringtoolbox.com/. [Accessed: 04-Oct-2018].

Supplementary Materials

Table A.1 illustrates the elemental analysis of the raw material used as feed for this study, while Table A.2 presents the proximate analysis of beech wood based on thermogravimetric measurements.

Table A.1: Elemental analysis of biomass used

Biomass/Biomass Component used	Elemental analysis (wt. %)			
	Carb-on	Hydrogen	Nitrogen	Oxygen
Beech wood (BW)	47.38	6.11	<0.01	46.51

Table A.2: Proximate analysis of different raw materials based on TGA experiments (wt. %)

	Humidity	Volatile matter	Fixed carbon	Ash
Beech wood (BW)	7.44	74.19	17.52	0.85

Hydrodynamic calculations for pyrolysis reaction occurring in DTR

- Determining terminal velocity, ut, of biomass particles

Using Stoke's law, $u_t = \sqrt{\frac{4 \cdot g \cdot d\,(\rho_{N_2} - \rho_{biomass})}{3\,\rho_{N_2} \cdot C_D}}$, for 0.1<Re<1000

Where ρ_{N_2} = 0.5956 kg/m^3 and $\rho_{biomass}$ = 840 kg/m^3

d: biomass particle diameter = 0.0004 m

g: gravity = 9.812 m/s^2 [25]

Re: Reynold's number = $\frac{\rho_{N_2} \cdot D \cdot u_t}{\mu}$,

where μ: dynamic viscosity of carrier gas = 3.49 × 10^{-5} kg/m.s [25]

C_D: drag coefficient = $\left(\frac{24}{Re}\right)(1 + 0.14Re^{0.7})$

To determine the terminal velocity of the biomass particle, an iterative calculation is required. The Reynold's number, Re_0 is first calculated for an initial value of $u_{t,0}$ = 3 m/s [21], followed by the drag coefficient $C_{D,0}$ and finally, once again, the terminal velocity, $u_{t,1}$. This process repeats itself; the values of Re_1 (using $u_{t,1}$), $C_{D,1}$ and then, $u_{t,2}$ and so on are obtained n times until $(u_{t,n}^2 - u_{t,n-1}^2)^2 = 0$.

The values thus found were: Re = 15.41; C_D = 3.04; u_t = 1.08 m/s

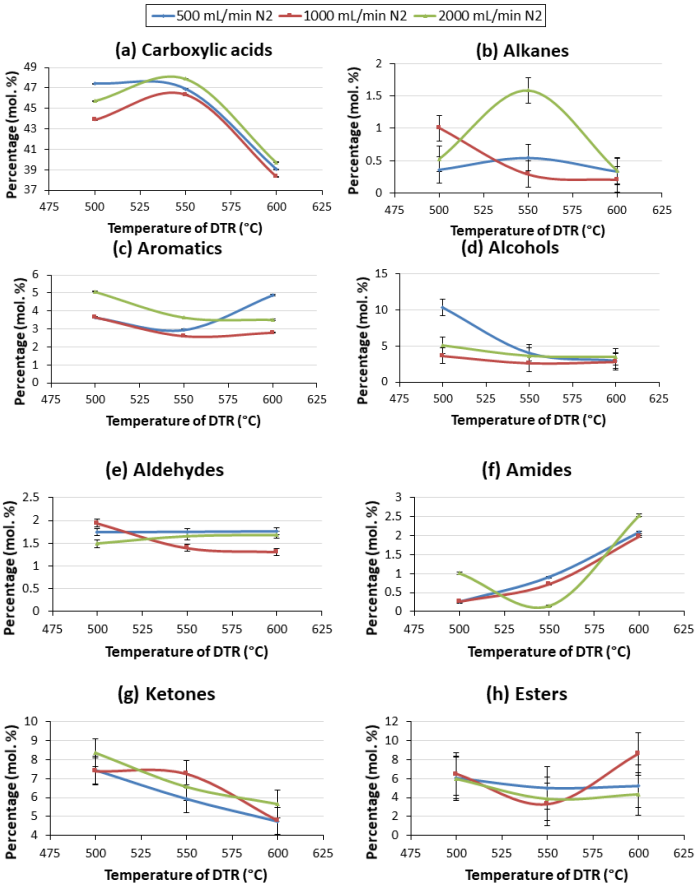
∴ $t_{particle}$: residence time of biomass particle = $\left(\frac{L_r}{u_t}\right)$ = 2.20 s, where L_r is the length of the reactor.

The following Table A.3 details the different NCG compositions obtained for the various experimental runs performed.

Table A.3: NCG compositions (vol. %) obtained for all experimental runs

Experiment performed	vol. %					
	CO	CO_2	CH_4	H_2	C_2	C_3
500 °C, 0.5 L/min N_2	58.27	15.86	12.46	5.39	4.60	3.43
500 °C, 1 L/min N_2	59.06	15.16	12.08	5.69	4.51	3.50
500 °C, 2 L/min N_2	59.80	15.53	11.86	5.23	4.42	3.15
550 °C, 0.5 L/min N_2	59.38	10.54	13.68	7.98	5.41	3.00
550 °C, 1 L/min N_2	60.08	10.47	13.27	7.84	5.30	3.04
550 °C, 2 L/min N_2	60.46	10.44	12.69	7.95	5.19	3.28
600 °C, 0.5 L/min N_2	56.32	9.70	15.64	10.14	6.00	2.19
600 °C, 1 L/min N_2	56.60	9.59	15.11	10.41	5.93	2.36
600 °C, 2 L/min N_2	56.44	9.56	14.56	11.17	5.79	2.48

Figures A.1 (a)-(l) illustrate the evolution of the different chemical families present in the bio-oil samples obtained under different nitrogen flow rates with increasing DTR temperature.



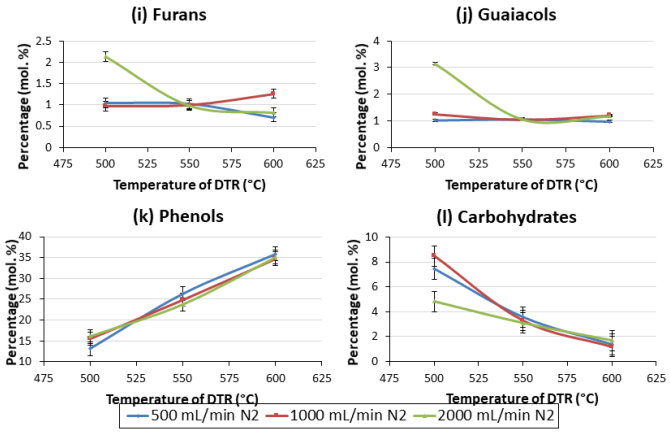


Figure A.1 (a)-(l): Evolution of chemical families (mol. %) in bio-oil samples obtained under different nitrogen flow rates with DTR temperature

Figures A.2 and A.3 depict the evolutions of the non-condensable gas compositions obtained with increasing DTR temperature under different gas flow rates, except that obtained with 500 mL/min N_2, as it was presented in the main text.

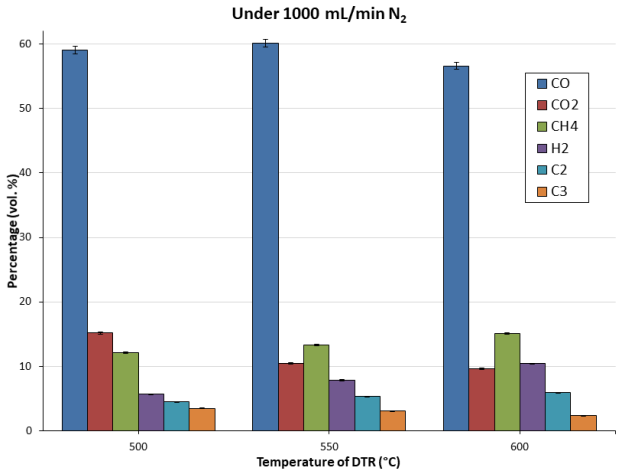


Figure A.2: Evolution of NCG components (vol. %) with DTR temperature at τ = 4.63 min

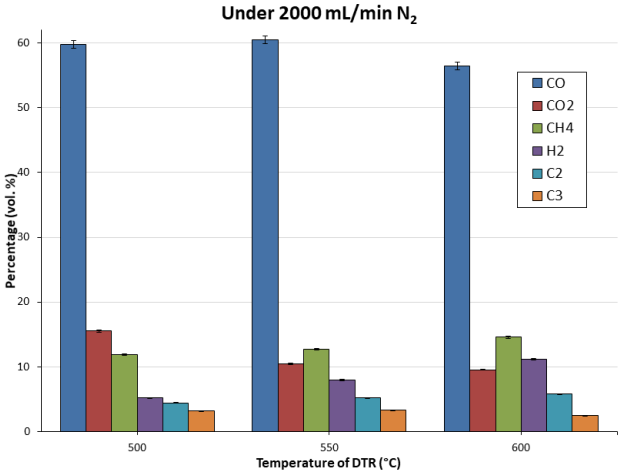


Figure A.3: Evolution of NCG components (vol. %) with DTR temperature at τ = 2.32 min

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 6, 443-447 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

CSFs for the Implementation of the Hybrid Lean ERP System. Stakeholders Interactions

Mariam Houti*, Laila El Abbadi, Abdellah Abouabdellah

Systems Engineering Laboratory, National School of Applied Sciences of Kenitra, Ibn Tofail University, PoBox 241, Morocco

ARTICLE INFO

Article history:
Received: 21 October, 2019
Accepted: 12 December, 2019
Online: 25 December, 2019

Keywords:
Lean ERP
CSFs
Performance
Stakeholders

ABSTRACT

Lean ERP system is considered among the new trends of systems in the manufacturing sector, it has become one of the most widely adopted systems, thanks to its various benefits such as: a centralized database which gathers all the data of the enterprise in a single database which is accessible to all the departments, reduction in costs, high performance and best quality and many other advantages. However, many companies are starting the implementation of the hybrid Lean ERP system without prior study of the project. This may lead to its improper implementation which can impact negatively the enterprise and its stakeholders.

It is in this context that our study aims to help future companies wishing to implement the hybrid Lean ERP system, through identification, study and an analysis of critical success factors (CSFs) of Lean ERP necessary to this implementation, while focusing on different groups of stakeholders and their interactions with CSFs.

1. Introduction

As businesses in manufacturing and service sectors are increasingly subject to global competition, the highly variable demand from customers and the growing variety of products on the market [1], enterprises face several challenges in order to maintain their places in the market, which requires the improvement of their performance and competitiveness [2,3]. Otherwise, it encourages them to increase productivity and product quality, reduce costs and cycle times, be more flexible and respect delivery deadlines [4].

In order to cope with these fluid conditions, manufacturers often use two strategies to face these challenges: implementing Lean practices and integrating information and communication systems, in particular ERP systems "Enterprise Resource Planning" [3]. This combined implementation gave birth to the new hybrid Lean ERP system, that combines the entire database of all department in a centralized one.

A good implementation of the hybrid Lean ERP system offers several benefits to the enterprise and its stakeholders through a significant reduction in costs, operating time and a considerable increase in performance and quality (which could be traced from raw material to finished good) which leads to: responding quickly to customer's needs. However, if the implementation is not well achieved, it could lead to serious problems and losses. It is therefore important to take into account the different parts of the company and critical success factors (CSFs) to ensure the successful implementation of the hybrid system.

The purpose of this Research is to identify Critical Success Factors (CSFs) for the implementation of the hybrid Lean ERP system, then study the degree of influence of the CSFs in order to extract the most decisive factors during implementation. Afterward, explore the role of stakeholders in the implementation of the hybrid system, in order to categorize which factors directly influence them, for the purpose of a successful implementation within the enterprise and its stakeholders.

2. CSFs For Lean ERP System

Critical success factors are all elements to consider when setting up or upgrading a new system, to optimize resources and effort by focusing on the important elements that contribute to successful implementation or upgrade. As a result, many researchers had shed more light on those factors (CSFs) and defined them according to their own points of view; the reason why we find different definition to CSFs in the literature.

Critical success factors usually vary from one industry to another, they define the elements influenced by the decisions of the top management and could affect the global competitive position of diverse sectors of an industry [5]. CSFs tend to relate more to how an organization approaches the change effort versus

*Corresponding Author: Mariam Houti | Email: Mariam.houti@gmail.com

change method specific factors. Even though certain authors bring up key aspects for the specific change method, most CSF are fairly general in nature [6]. Houti concluded that the concept of critical success factors is one of the important elements that influence the success of any organization. It should be used when implementing a system. As they are among the few elements that guarantee successful implementation in all phases: pre-implementation, implementation and post-implementation. While ensuring the sustainability of the system, through the use of factors of measurement of profitability and performance, as well as the respect of the different characteristics and conditions imposed in the specifications of the business [4]. Knowing that CSFs can have a direct positive or negative effect on the performance of the business. In fact, if they are well identifiable and applied, they can lead to the achievement of the business. On the other hand, if they are not used accurately they can lead to failure [4].

Several authors have discussed the critical factors whether success or Failure. Researchers have already identified according to the literature that CSFs affecting the implementation of both production systems, namely: Lean Manufacturing and ERP system. Our study is based on the work established by Houti and al., [4, 2] to gather the most critical success factors of Lean ERP system. It will be necessary to identify factors of success to guarantee an effective implementation of the combined system.

Our study allowed us to identify 20 different critical factors of success during the Implementation of the combined Lean ERP system. They cover the categories of management, resources and organizations, as well as the phases of implementation the Lean ERP system. The factors gathered are listed in the table 1:

Table 1. CSFs for Lean ERP implementation

N°	CSFs
C 1.	Top management support and commitment
C 2.	Project management and planning
C 3.	Change management / Organizational culture / Political issues;
C 4.	Business plan and a clear vision;
C 5.	Training and education;
C 6.	Effective Interdepartmental communication;
C 7.	Dedicated resources;
C 8.	Adequate ERP software selection and good Lean tools;
C 9.	Project monitoring and evaluation of performance;
C 10.	Active, strategic and visionary leadership;
C 11.	Suitable IT legacy systems;
C 12.	Interdepartmental cooperation;
C 13.	Skills and expertise;
C 14.	Involvement of stakeholders;
C 15.	Management of expectations;
C 16.	Implementation Strategies;
C 17.	Business Process Reorganization (BPR);
C 18.	Project selection and prioritization;
C 19.	On-going support (Customer, vendor);
C 20.	Motivation;

2.1. Degree of Influence Of Csfs For Lean ERP Implementation

These 20 CSFs identified above have several interactions with each other. Our study is established thanks to an effective collaboration with a team of researchers and industrialists, in order to extract the different relations existing between each of these factors, and which are visualized later using the vectograph drawing method, represented in FIG. 1, where:

- The "X" represents the factors on the columns influence the factors on the rows.
- The "O" represents the factors on the rows influence the factors on the columns;

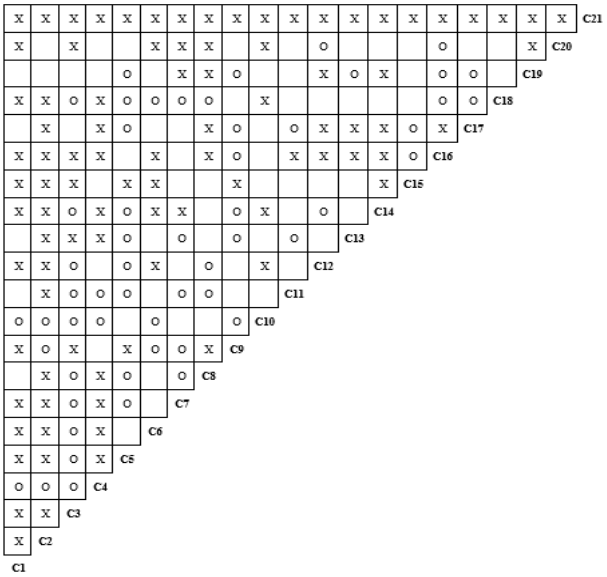


Figure 1. The influence among the different CSFs in Lean ERP implementation

The analysis of the figure 1 above, allowed us afterwards to draw the binary matrix [M] which is illustrated in figure 2, to better visualize the influence of each factor separately, Then, through calculating the degree of influence of each factor, the result matrix could be plotted, where:1 represents an influence and 0 doesn't represent any influence;

By analyzing the resulting matrix [M], we can conclude that there are 11 levels of influence, distributed as follows in Table 2:

These CSFs's levels and degrees of influence will contribute to an easy implementation of the Lean ERP combined system; Starting with the first levels of influence (high influence factors), it will contribute indirectly to the establishment of other factors influenced factors; This will automatically facilitate the implementation of the other CSFs, and thus a quick implementation of the hole system.

3. Stakeholders Role in Lean ERP Implementation

Stakeholders could be defined as the different parts involved in a process, hence their positive or negative impact; depending on their degree of contribution and influence on the project. Knowing that close cooperation of stakeholders is necessary for the success of the project [7].

Identifying how stakeholders influence the success of project through a continuous development of their relationships, is an important and fundamental issue of stakeholder management, that remain an account for the success of projects [8], and an important issue in project management [9]. It is necessary to know how the stakeholders can be influenced, so they support and contribute to the project [10]; As a project can be seen as a temporary coalition of stakeholders pursuing an aim together [7,9].

$$M = \begin{array}{c} \\ C1 \\ C2 \\ C3 \\ C4 \\ C5 \\ C6 \\ C7 \\ C8 \\ C9 \\ C10 \\ C11 \\ C12 \\ C13 \\ C14 \\ C15 \\ C16 \\ C17 \\ C18 \\ C19 \\ C20 \end{array} \begin{array}{c} \begin{array}{cccccccccccccccccccccc} C1 & C2 & C3 & C4 & C5 & C6 & C7 & C8 & C9 & C10 & C11 & C12 & C13 & C14 & C15 & C16 & C17 & C18 & C19 & C20 & C21 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array} = \begin{Bmatrix} 13 \\ 14 \\ 6 \\ 13 \\ 4 \\ 7 \\ 6 \\ 8 \\ 5 \\ 11 \\ 7 \\ 6 \\ 8 \\ 9 \\ 1 \\ 4 \\ 5 \\ 8 \\ 7 \\ 3 \end{Bmatrix}$$

Table 2. CSFs's Degree of influence

Level	Degree Of Influence	Factors
Level 1	14	C2
Level 2	13	C1 C4
Level 3	11	C10
Level 4	9	C14
Level 5	8	C8 C18 C13
Level 6	7	C11 C19 C6
Level 7	6	C12 C3 C7
Level 8	5	C17 C9
Level 9	4	C16 C5
Level 10	3	C20
Level 11	1	C15

The implementation of Lean ERP system is one of the complex project that need contribution and engagement of all part of the enterprises, due to the changes that will involve stakeholders. These stakeholders typically have multiple and often conflicting interests and rarely agree on a set of common objectives [11].

3.1. Classification according to Stakeholders Group

Stakeholder involvement is essential when implementing a system such as Lean ERP; That provides high customer service and shorter lead times [12], which involves different parts of the enterprises. why is it important to consider the different categories of stakeholders to allow each group to focus on the relevant CSFs. Figure 3 shows the important groups of stakeholders. The Lean ERP implementation project involves all parts of the business. Therefore, the assessment of stakeholder groups is one of the important steps to establish when implementing the combined system.

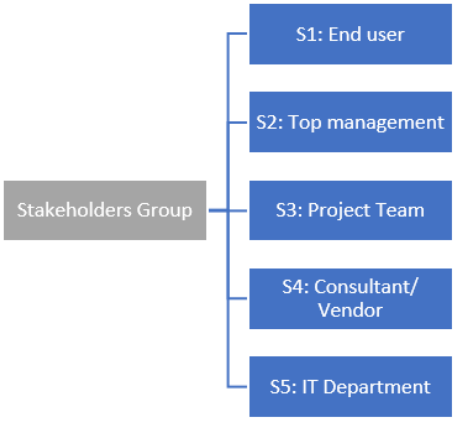


Figure 3. Classification of stakeholders groups

3.2. Stakeholder Assessments

Stakeholder assessments can assist planners to identify the interests, groups and individuals that are stakeholders in planning exercises, as well as their values, interests and relative power [13]. This can help to proactively determine ways to reduce negative impacts on the groups with less influence and power within the enterprise. A thorough stakeholder analysis can also identify potential conflicts or risks that could jeopardize the project, as well

Table 3. Analysis of Stakeholder assessments

Stakeholders	Role	Influence	Expectation	Issues
End user	▪ Provide all the necessary business processes ▪ Effective communication of IT needs of the enterprise and its consultants	Low influence	▪ solved Business process ▪ Centralized system offering real and accurate information	▪ Lack of involvement in decision-making
Top Management	▪ Provide the resources needed to implement the project ▪ Set up a road map and an execution plan ▪ Ensure stakeholders involvement	Medium influence	▪ Successful implementation of the combined system ▪ Minimize waste, reduce costs to increase profits	▪ Lack of support and ongoing commitment ▪ Implementation of other projects in parallel that minimize starting resources
Project Team	▪ Application of business plans developed by Top Management ▪ Good use of available resources ▪ Adopt lean philosophy throughout the implementation of the Lean ERP system ▪ Ensure stakeholders involvement	High influence	▪ Accurate monitoring of the plans and processes put in place ▪ Provide relevant employees with the necessary information and documentation in real time, with precision.	▪ Anxious to be replaced by the new system
Consultant / Vendor	▪ Work methodically with internal stakeholders ▪ Provide all the necessary documentation and training ▪ Ensure the successful implementation of the system	High influence	▪ Successful and efficient implementation of the combined system ▪ Collaboration and support of internal stakeholders	▪ Lack of cooperation from internal stakeholders
IT Department	▪ Technical support ▪ Collaboration with consultant/ Vendor and Project Team to guarantee an effective implementation of the combined system	Low influence	▪ Consolidation of decentralized IT groups ▪ Creation of a centralized IT department	▪ Lack of cooperation ▪ Lack of involvement in decision making process

as opportunities and strategies for stakeholder engagement during implementation. Therefore, it is imperative to identify and engage all stakeholders early on in the project to ensure the proper measures are in place to manage the different characteristics of these stakeholders. Taking this initiative will increase the likelihood of successful implementation and adoption of the enterprise system [14]. Table 3 summarizes the assessment of stakeholder's groups including their role, degree of influence, expectations and issues related to Lean ERP implementation.

The stakeholder groups mentioned above have multiple interactions with each other within the organization, which is essential to the smooth running of the business, which has prompted our team of researchers and industrialists to draw the matrix [S](Figure 4) which standardize the degree of similarity and interactions over an interval of 0 to 1.

$$S = \begin{array}{c} \\ S1 \\ S2 \\ S3 \\ S4 \\ S5 \end{array} \begin{array}{c} \begin{array}{ccccc} S1 & S2 & S3 & S4 & S5 \end{array} \\ \begin{pmatrix} 1 & & & & \\ 0.9 & 1 & & & \\ 1 & 0.9 & 1 & & \\ 0.4 & 0.7 & 0.8 & 1 & \\ 0.9 & 0.5 & 0.9 & 0.7 & 1 \end{pmatrix} \end{array}$$

Figure 4. The binary Matrix [S]

Afterwards, we used the color-coding approach in figure 5to illustrate the different interactions of the stakeholder network. The color ranges from turquoise to gray to indicate the degree of interaction of the stakeholder network. If the links colors are closer to gray then we have a higher degree, which means that related stakeholders have a closer connection and stronger influence over each other. On the other hand, the colors closer to turquoise are not very marked, which means that the related stakeholders have a fragile relationship and a weaker influence on each other. Table 4 presents this classification that will help stakeholders of the enterprise to focus on factors that affect them.

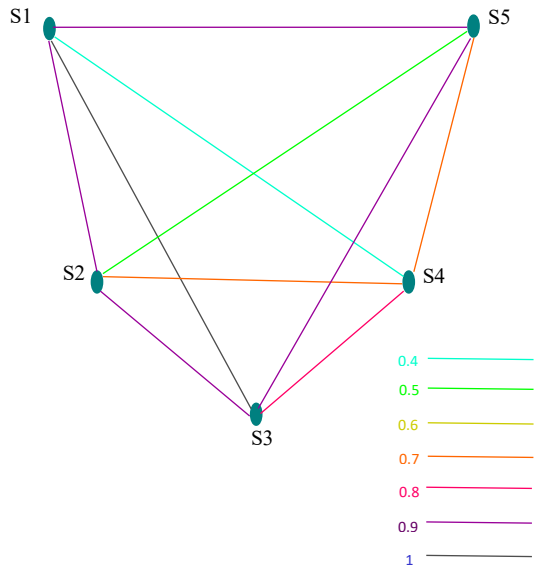


Figure 5.Relationships among Stakeholders network

Consequently, a classification according to Stakeholders Group of CSFs gathered during the previous study will be established to help further enterprises and its stakeholders to focus on CSFs that affect them directly in order to reach a better performance and a successful implementation of the hybrid Lean ERP system.

Table 4. Classification of Lean ERP CSFs according to stakeholders group

CSF \ Stakeholders Group	S1	S2	S3	S4	S5
C1.		X			
C2.		X	X		
C3.		X	X		
C4.		X			
C5.	X		X		X
C6.	X		X		
C7.		X			X
C8.		X		X	X
C9.			X		
C10.		X			
C11.					X
C12.	X		X		
C13.			X	X	X
C14.	X				
C15.		X	X	X	
C16.		X			X
C17.	X	X		X	
C18.		X		X	
C19.	X			X	
C20.		X	X		

4. Conclusion

The Lean ERP system has become the most widely adopted system by enterprises, as long as the many advantages that presents. However, the high cost of implementation and the risk of failure can lead to several critical issues within the enterprise and its different stakeholders, which could be solved by taking into account the CSFS and involving different stakeholders groups during the implementation of the hybrid system.

This article has enabled us to extract, 20 most CSFs to be used when implementing the hybrid Lean ERP system, thereafter study in more detail their degree of influence to further identify the factors to take into account as a priority, depending on their Level of influence, in order to simplify the implementation of other factors that affect them. Thereafter analyze the importance of the identified stakeholder groups (End user, Top Management, Project Team, Consultant / Vendor, IT Department) and study their interactions to help future enterprises wishing to implement the hybrid system and their stakeholders to concentrate on the CSFs that directly affect each Stakeholder group in order to achieve better performance and successful implementation of the hybrid Lean ERP system.

5. References

[1] Houti, M., El Abbadi, L., Abouabdellah, A., "E-Kanban the new generation of traditional Kanban system, and the impact of its implementation in the enterprise ", Proceedings of the International Conference on Industrial Engineering and Operations Management, IEOM., Rabat, Morocco, (Apr, 11-13), 2017.

[2] Houti, M., El Abbadi, L., Abouabdellah, A., "External critical successful factors for a successful implementation of ERP systems", Proceedings of International Conference on Computers and Industrial Engineering, CIE., Lisbon, Portugal, (Oct, 11-13), 2017.

[3] Houti, M., El Abbadi, L., Abouabdellah, A., "Lean ERP: A hybrid approach Push /Pull", Proceedings of the 3rd IEEE International Conference on Logistics Operations Management, GOL 2016., Fez, Morocco, (May, 23-25), 2016.

[4] Houti, M., El Abbadi, L., Abouabdellah, A., "Critical Success Factors for Lean implementation "Projection on SMEs", Proceedings of the International Conference on Industrial Engineering and Operations Management, IEOM., Pilsen, Czech Republic, (Jul, 23-26), 2019.

[5] Hofer, C. W., Schendel, D. A., "Strategy Formulation: Analytical Concepts", West Publishing Company, Saint Paul, 1978.

[6] Dag, N., "Lean and Six Sigma – critical success factors revisited", *International Journal of Quality and Service Sciences*, Vol. 5, N°.1, (2013), pp. 86 - 100.

[7] Jepsen, A.L., Eskerod, P., "Stakeholder analysis in projects: Challenges in using current guidelines in the real world", *International Journal of Project Management,* Vol. 27, (2009), pp. 335–343.

[8] McElroy, B.; Mills, C. "Managing stakeholders", In Gower Handbook of Project Management; Turner, R.J., Simister, S.J., Eds.; Gower: Aldershot, UK, 2000.

[9] Andersen, ES., "Prosjektledelse – et organisasjonsperspektiv", Bekkestua: NKI Forlaget, Oslo, 2005.

[10] Lund Jepsen, A., Eskerod, P., "Stakeholder analysis in projects: Challenges in using current guidelines in the real world", *International Journal of Project Management*, Vol. 27, (2009), pp. 335–343.

[11] Sudevan, S., Bhasi, M., Pramod, K.V., "Interpreting Stakeholder Roles in ERP Implementation Projects: A Case Study", *International Journal of Computer Science and Information Technologies,* Vol. 5 (3), (2014), pp. 3011-3018.

[12] Houti, M., El Abbadi, L., Abouabdellah, A., "How Could an Urgent Order Disturb the Supply Chain? Case S tudy of an Automotive Industry", *International Journal of Engineering & Technology*, Vol. 7(4.16), (2018), pp. 155-159.

[13] Medeiros de Araujo. L., Bramwell. B., "Stakeholder Assessment and Collaborative Tourism Planning: The Case of Brazil's Costa Dourada Project", *Journal of Sustainable Tourism*, Vol. 7(3-4), (1999), pp.356–378.

[14] McLaren, T., Jariri, I., "Stakeholder Assessment and Management for Enterprise Systems Implementation Projects", International Conference on Information Resources Management, CONF-IRM 2012 Proceedings, Association for Information SystemsAIS Electronic Library (AISeL) (sep, 26), 2012.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 452-459 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Promoting Continuous Professional Development Among Academics from A Vocational College by Using A Practical Workshop Based on Arduino Technology

Arthur James Swart*, Pierre Eduard Hertzog

Central University of Technology, Faculty of Engineering, Built Environment and Information Technology, Bloemfontein, South Africa

ARTICLE INFO

Article history:
Received: 17 November, 2019
Accepted: 27 March, 2020
Online: 04 April, 2020

Keywords:
practice
e-learning
pre-post test

ABSTRACT

The internet of things has resulted in the design and development of many electronic sensors that need to be integrated with various technologies. Engineering academics need ongoing training relating to this integration as they need to subsequently teach their students. This relates to continuous professional development. The purpose of this article is to illustrate how engineering academics from a vocational college have been able to engage with the theory and practice of sensor integration with Arduino technology, thereby strengthening or reconstructing their conceptual knowledge of these electronic components. A case study is used where experts in electrical engineering designed and facilitated an academic development workshop (2 days in duration) to address a need for continuous professional development. Eighteen academics from a local vocational college attended this workshop where the first goal was to determine their conceptual knowledge of 10 electronic components using an electronic responsive system (pre-test). The second goal was to facilitate a "hands-on" laboratory session where academics had to physically integrate various sensors with an Arduino microcontroller. A post-test online questionnaire was then used to determine their conceptual knowledge again, in order to ascertain the impact of the workshop. Results indicate that their conceptual knowledge was reconstructed with regard to the purpose and use of capacitors, Zener diodes, transformers, ultrasonic sensors, reed switches and passive infrared receivers. A problematic question arose relating to photocells in the first workshop that was held in 2017, which was resolved by the facilitators prior to offering a second workshop in 2018. Positive feedback was received from the participants regarding the relevance and presentation of the workshops. The novelty of this work relates to the design of a practical workshop where all participants are engaged to either strengthen, or reconstruct, their conceptual knowledge regarding electronic components that may be integrated with Arduino technology.

1. Introduction

This article is an extension of work originally presented at the 12th International Conference on Sensing Technologies, held in Limerick Ireland from 3-6 December 2018. The original paper considered the importance of academic development workshops and focussed on the results of one such workshop that was offered to academics from a vocational college in 2017 [1]. The current article adds to the discussion by highlighting the importance of continuous professional development (as suggested by the reviewers), by detailing the four practical assignments that were designed for the workshop and by including the results of a second workshop that was offered in 2018. The results of this second workshop are especially important, as a question relating to one of the sensor components caused some confusion among the participants in the first workshop offered in 2017. The revised question attained the goal of strengthening the conceptual knowledge of the participants.

Students or academics need to be engaged on a personal level, where their conceptual knowledge is either strengthened or reconstructed in order to enhance their intellectual capacity, thereby enriching their lives. Reconstructing knowledge refers to identifying gaps or errors in one's thinking, which first needs to be

* Arthur James Swart, 685 Email : drjamesswart@gmail.com

deconstructed before it can be reconstructed again in the correct way. One way of achieving this is by fusing theory and practice, which is mandatory in engineering education [2, 3]. Students must be allowed to experiment actively in a laboratory, where theory and practice are linked to real-world problems. Academics must be allowed to put into practice their acquired theoretical knowledge (new or old) in an academic development workshop that could lead them to enhancing their teaching practice.

Academic development workshops are successful tools of change in teaching practice and post-workshop evaluations operate as indicators of impact [4]. An academic's teaching practice must be investigated on a regular basis to determine areas of self-improvement. This forms part of the Scholarship of Teaching and Learning [5] and also of continuous professional development [6]. TVET (technical and vocational education and training) college academics should also engage with this, as they seek to prepare their students to meet the needs of industry, and especially as engineering technicians. Their conceptual knowledge of their field should be strengthened, or reconstructed, on a regular basis, so as to build capacity within the college regarding new technologies. Such expertise is required within any discipline, including electrical engineering, where the study and application of electronic components (including sensors) and microcontrollers is constantly evolving.

Many microcontrollers require a software program to execute commands [7] that can be achieved by significant "hands-on" practice [8]. The internet of things has been defined as a massive network where a diversity of electronic components are used to collect all kinds of data, including sound, light, heat, electricity, mechanics, chemistry, biology, and position [9]. Training academics to effectively integrate many of these electronic components with a specific microcontroller may enable them to teach their students more effectively, and thereby helping academics fulfil their teaching responsibilities. It may further help students to be better equipped to meet the needs of Industry.

The purpose of this research is to illustrate how engineering academics from a vocational college have been able to engage with the theory and practice of sensor integration with Arduino technology, thereby strengthening or reconstructing their conceptual knowledge of these electronic components. The link between academic development workshops and continuous professional development is firstly explained, along with the structure of the practical workshop. The research methodology, results and conclusions then follow.

2. Academic Development Workshops Linked to Continuous Professional Development

Research suggests that relevant academic development workshops are critical to enhance the quality of education, as they enable participants to acquire new and innovative teaching and learning practices [10]. Education should never be static. This requires academics to constantly reflect on their teaching practices, seeking innovative ways to improve it. Academic development workshops provide opportunities for such improvement by promoting a culture of reflective practice [11]. These workshops are vital within a dynamic higher education landscape. However, to be effective, they need to be facilitated by experts within a relevant field of study [12]. They should also aim to fuse theory and practice, which may be achieved through a practical workshop. Goals of such workshops must include the transfer of theoretical and technological knowledge, the acquisition of practical hands-on skills [13] and the reinforcement of conceptual knowledge through practice [1].

A practical workshop regarding the production of Educational videos was held at the University of Huddersfield. This was a '*hands-on'* workshop where the participants worked in pairs developing a script and a visual storyboard for their educational videos. Participants learnt about the range of videos and multimedia that can be used in the educational environment and how this could *benefit learners* [14].

A practical workshop on engineering ethics was held at the University of Nottingham in Malaysia. A post-workshop survey revealed that the participants knowledge on engineering ethics had not significantly improved, but it had *motivated* them to act more ethically [15].

A further important goal of practical workshops should be to address Continuing Professional Development (CPD). CPD has been defined as the maintenance and enhancement of knowledge expertise and competence of professionals throughout their careers according to a plan that is formulated to meet the needs of an individual [16]. In a PhD study that explored the impact of short CPD courses on the professional development of teachers, it was found that only one CPD course (duration of one day) focusing on subject-specific information had valuable impacts on the teacher's practice [17].

Demonstration of CPD is increasingly required for re-certification of professionals in the workplace [18]. The certification of professional engineers in SA is done by the Engineering Council of South Africa (ECSA). ECSA has an online CPD recording system that consists of 3 main categories, namely development activities, work-based activities and individual activities. Development activities include attending conferences, workshops and colloquiums. Work-based activities include engineering work and mentoring candidate engineering practitioners. Individual activities include supervision of postgraduate students, evaluation of dissertations, publishing research articles and membership of a recognized voluntary association, such as the IEEE. A minimum of 25 credits must be obtained over a five-year cycle to remain registered as a professional Engineer or Technologist. Only ECSA recognised voluntary associations and tertiary institutions with accredited engineering programs may validate CPD developmental activities [19].

3. Structuring Practical Workshops

Practical workshops should be bottom-up approaches. They should provide equal engagement by all participants [20]. This can be achieved by using an electronic responsive system (ERS) where individual participant engagement is maintained. Attending a workshop does not guarantee meaningful engagement, as the participant may not actively engage in the discussions due to shyness or distractions. Encouraging active engagement is required to achieve the following primary outcomes of any developmental workshop:

- Agency; individuals need to act voluntarily and independently, making their own decisions [21].

- Ownership; individuals need to actively participate in decision taking [22].
- Active learning; individuals need to be personally involved [23] by pressing knobs, sketching images, etc.
- Satisfaction of learning; individuals learns from their own misconceptions, as their ideas are shared and scrutinized by others in a "secretive" way [24].

In this research, training is defined as strengthening, or reconstructing, the conceptual knowledge of TVET academics regarding the integration of electronic components with Arduino technology. An academic development workshop, or practical workshop, is used to achieve this aim.

4. The Practical Workshop Used in This Research

The practical workshop reported on in this article featured two stages. In stage one, the conceptual knowledge of the participants was reviewed with regard to 10 electronic components (including 5 sensors). In stage two, the software programming skills of the participants were enhanced through repeated practice and interaction with an Arduino microcontroller. The practical workshop was designed by experts in electrical engineering who have published extensively on this topic. Their most recent contributions considered the effective use of Arduino in engineering education [25], the use of Arduino in energy monitoring [26] and the use of Arduino in evaluating the performance of pico-solar systems [27].

The first stage of the workshop required participants to engage with two Microsoft PowerPoint presentations. The main presentation covered the workshop program, learning outcomes and requirements for CPD. Content relating to the purpose and application of 10 electronic components and sensors were also included. A second presentation was designed using Turning Point Software that enables interactivity between the facilitator and participants. This presentation contained a number of multiple-choice questions that were answered in real-time by the participants, in an effort to determine their conceptual knowledge of the selected components and sensors.

This interactive presentation necessitates the use of a transmitter (resembling a calculator) and receiver (USB device connected to a PC). Voluntary responses of each participant (*agency* involved) is transmitted (*active learning* as keys need to be pressed) wirelessly to the receiver. Once all responses have been received, then the results are visually displayed on the next slide for all to see and analyse (this contributes to *ownership* as each contributes to the discussion). Discussions on the real-time results then follows to help the participants to either strengthen their conceptual knowledge of the components and sensors, or to deconstruct and then reconstruct it again (*satisfaction of learning* occurs as perceptions are shared and compared). These results also enable the facilitator to identify gaps or misconceptions in the thinking of the participants, thereby spending more time clarifying the purpose and application of specific components. The results of this interactive presentation were saved for future analysis.

In the second stage of the practical workshop, the facilitators discussed the basic fundamental theory of the hardware and software associated with the Arduino microprocessor. This included:

- Identifying the key differences between the UNO and MEGA microprocessors;
- Listing the available inputs, outputs and basic connections on the UNO microprocessor;
- Structuring a basic C program in the Arduino integrated development environment (IDE); and
- Reviewing various codes for integrating components and sensors onto the UNO microprocessor.

Four practical assignments were designed for the participants to complete, with each one requiring a specific software program that had to be uploaded to an UNO microprocessor. Four steps had to be followed for each assignment

- Compile a basic software program on the computer;
- Wire the component or sensor to the microprocessor using a breadboard;
- Upload the program using a USB cable; and
- Troubleshoot any errors.

The first assignment required the participants to connect a Liquid Crystal Display (LCD) to the UNO microprocessor where the name of the participant had to be visually shown on the display. The second assignment required the participants to integrate a light dependent resistor (LDR), or photocell, with the UNO microprocessor, where the previously used LCD would now display the resistance value of the LDR. The third assignment required the participants to integrate with the Arduino microprocessor to measure distance and then display it on the LCD. This assignment is described next in detail with regard to the four steps that participants needed to follow. In this way, each assignment built one the first one, as the LCD was subsequently used in all the remaining assignments. This is a form of scaffolding, where support was initially given by the facilitators to the participants regarding the integration of the LCD in the first assignment. However, this support relating to the LCD was withdrawn from subsequent assignments and replaced with support for other components and sensors.

The **first step** requires a *software program code*. An example of this code is displayed in Figure 1. In line 1, the LCD library is included and in line 2 the connection pins for the LCD module is defined. The LCD is also defined as a two row LCD with 16 characters per row. The trigger pin of the ultrasonic sensor is also declared as an output and the eco pin as an input. In line 13 of the main loop, the ultrasonic function is called. The measured distance is calculated by this function and placed in the "distance" variable. The cursor position of the LCD is set to row 0 and character 0 where after "Ultrasonic" is written to the LCD. The cursor is then moved to character 0 of the second line and "distance:" is written to the position. The measured distance is then displayed on the LCD, with mm as the unit.

```
#include <LiquidCrystal.h>
LiquidCrystal lcd(12,11,5,4,3,2);
int trigPin = 9;
int echoPin = 10;
long distance;
long duration;
void setup(){
  lcd.begin(16, 2);
  pinMode(trigPin, OUTPUT);
  pinMode(echoPin, INPUT);
}
void loop(){
  ultrasonic();
  lcd.setCursor(0,0);
  lcd.clear();
  lcd.print("Ultrasonic");
  lcd.setCursor(0,1);
  lcd.print("distance : ");
  lcd.print(distance);
  lcd.print("cm");
  delay(200);
}
void ultrasonic(){
  digitalWrite(trigPin, LOW);
  delayMicroseconds(2);
  digitalWrite(trigPin, HIGH);
  delayMicroseconds(10);
  digitalWrite(trigPin, LOW);
  duration = pulseIn(echoPin, HIGH);
  distance = duration*0.034/2;
}
```

Figure 1: Arduino code for assignment 3

Figure 2 shows the schematic diagram for assignment 3. The participants use the information in this circuit diagram to physically integrate the required sensors with the Arduino microcontroller with the aid of a breadboard. The Arduino UNO, the LCD and the ultrasonic sensor can be observed in Figure 3 that represents the **second step** in the assignment, being the *wiring of the sensor* to the microcontroller. This wiring diagram provides a real-life image of each component, thereby helping the participant to select the correct component from the electronic kit that was provided to them in the laboratory. This diagram is drawn by the participants using a free electronic CAD software, named Fritzing that can be downloaded from the internet. This wiring diagram further assists the participants with fault-finding, as connection mistakes are common. This links to the **fourth step** of the assignment, namely *troubleshooting*, that is also required during the third step of the other three assignments.

Step three requires the participants to *upload the program* to the Arduino microprocessor via the IDE. The participants first need to resolve any errors, as indicated by the IDE (the line in the code that contains the error is highlighted). Common errors that novice programmers make may include format and punctuation in the commands. The commands are also case sensitive, and many errors stem from upper- and lower-case letters. Before uploading, the participants also need to set the output in the IDE to Arduino UNO if the UNO board is used, as there are many boards in the Arduino family. The output port in the IDE also needs to be set to the port that the Arduino UNO is connected to. Many upload errors stem from either wrong board or port selections before uploading the code. The upload process usually takes only a few seconds and, if successful, the program will immediately start executing on the Arduino microprocessor.

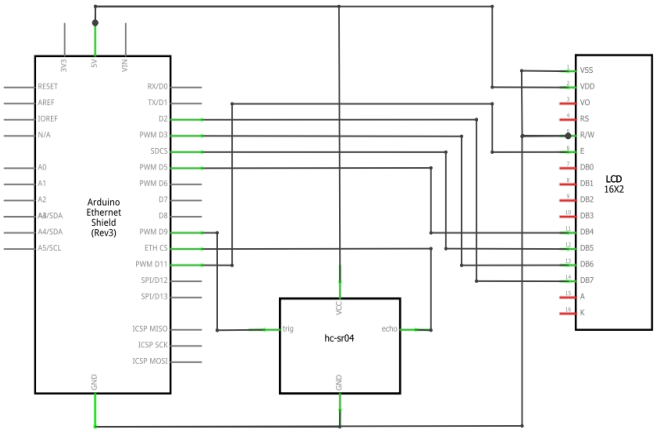


Figure 2: Schematic for assignment 3

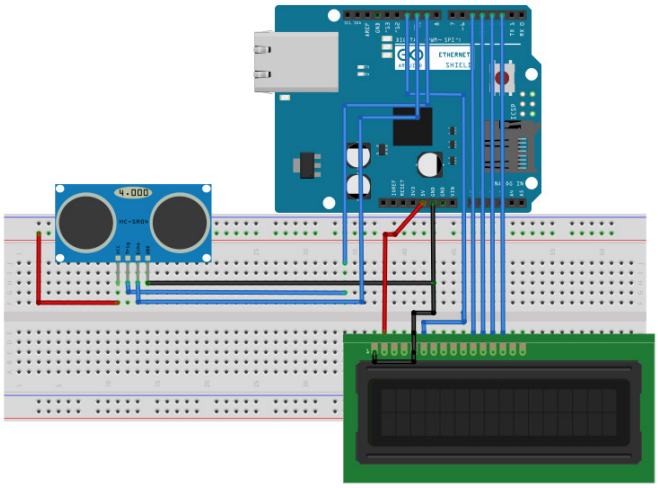


Figure 3: Breadboard for assignment 3

The learning outcome of the fourth assignment required a passive infrared sensor (PIR) and buzzer to be integrated with an UNO microprocessor, where the LCD would indicate an alarm condition. Participants had to again repeat the aforementioned steps for this assignment.

5. Research Methodology

A case study is used for this research where the results cannot be generalized. However, a case study can provide practical insights about industry-specific problems that are being addressed and therefore enable learning and changes in practice to occur [28]. In this research, insights into the conceptual knowledge of TVET academics are presented that can impact their workplace practice as they teach and train students in electrical engineering. This knowledge of the participants was either strengthened, or needed to be reconstructed, with regard to 10 electronic components and sensors. The findings of this case study further seek to establish if the designed workshop attained its aim of promoting CPD and if any changes needed to be made to enhance future workshops.

The practical workshop was first presented in 2017 that was attended by 10 TVET academics. The number of participants was limited to 10 to enable more personal attention to be given by the facilitators. This is especially required during the programming

and troubleshooting stages. A second workshop was presented in 2018 that was attended by 8 TVET academics.

A pre-post test, with no control group, was used to determine the impact of the practical workshop on the conceptual knowledge of the participants. A similar approach has been documented for nursing education [29]. Similar multiple-choice questions were posed for both tests that provided the participants with a predefined list of electronic components and sensors to choose from. Each question was phrased to cover a specific purpose relating to each component. The pre-test was administered during the first stage of the workshop using an ERS. The post-test was administered at the end of the second stage of the workshop using Google Forms. Grades from these tests were analysed quantitatively and are presented in the form of a radar chart. This helps to visually identify which electronic component the participants know well, and which ones require more discussion so that their conceptual knowledge may be reconstructed.

Feedback from the participants was also obtained in the post-test with regard to the facilitators and the workshop. Eight close-ended questions were asked using a 5-point Likert scale. The profile of the academics was initially obtained during the first stage of the workshop using the ERS. This also helped to train the participants in using the ERS correctly. The end goal of the profile is to better contextualise the results.

6. Results and Discussions

Figure 4 shows the profile of the workshop participants for 2017, while Figure 5 shows the profile for 2018. More than 90% of the participants were older than 30 years, validating them as academics with life experience. The noted disciplines of the academics validate the relevance of this Arduino based practical workshop, as the majority of the participants (60% + 50% from 2017 and 2018 respectively) were from electrical engineering. It is true that Arduino technology is used by individuals from a wide range of fields, including electrical engineering [25, 30]. The third part of the profile presents the home languages of the participants (right hand side of Figure 4 and Figure 5). Four languages are shown that is not uncommon for SA where 11 official languages are spoken [31]. A dominant language in 2017 was Afrikaans, while in 2018 it was Sesotho. Figure 6 illustrates the perceptions of the participants regarding the workshop facilitators.

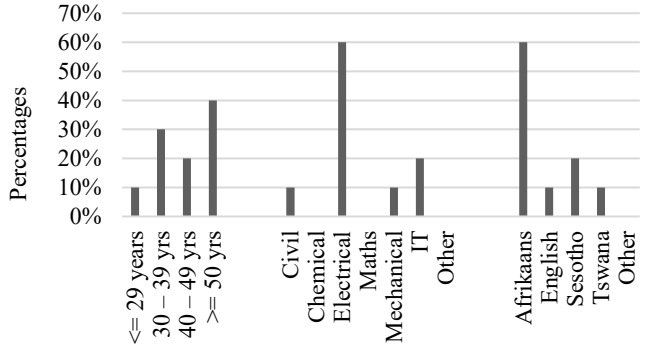


Figure 4: 2017 participant profile

The majority of the participants (67%) rated their time management skills as very good, while 89% rated their presentation skills in the same way. 89% of the participants agreed that the subject knowledge of the facilitators was very good. Finally, 94% of the participants rated the academic professionalism of the facilitators as very good. Literature suggests that good time management in education is vital [32] as are good presentation skills [33]. It is equally important that a facilitator personally possesses the relevant subject knowledge [12] and academic professionalism (this may include shared values, unselfish concern for others, relevant expertise, logical thinking, use of evidence, conceptual and theoretical rigour and the unbiased pursuit of truth [34].

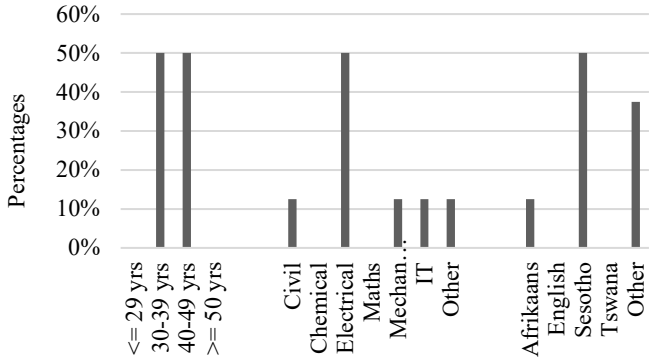


Figure 5: 2018 participant profile



Figure 6: Participant perceptions regarding the facilitators

Figure 7 indicates the perceptions of the participants with regard to the workshop itself. The majority of the participants (94%) agreed that the workshop venue was appropriate. Appropriate facilities enhance not only the learning process, but also contributes to student satisfaction [35]. Time management is of great importance to academics in TVET colleges who need to cover specific course content in a limited amount of time [36]. Wasting time is therefore not an option. Many of the participants expressed the view that the practical workshop was not a waste of time (83% disagreed with the given statement). The majority (61%) further agreed that the workshop was relevant to their work responsibilities. This correlates well with the disciplines shown in Figure 4 and Figure 5, where 10 out of the 18 participants (56%) came from electrical engineering. However, it further correlates with literature showing the relevance of Arduino in different disciplines [25, 30]. A good indication that the workshop was valuable relates to the fact that 89% of the participants strongly

agreed that they would recommend it to their colleagues. Table 1 shows the conceptual knowledge questions that were used in the pre-post test.

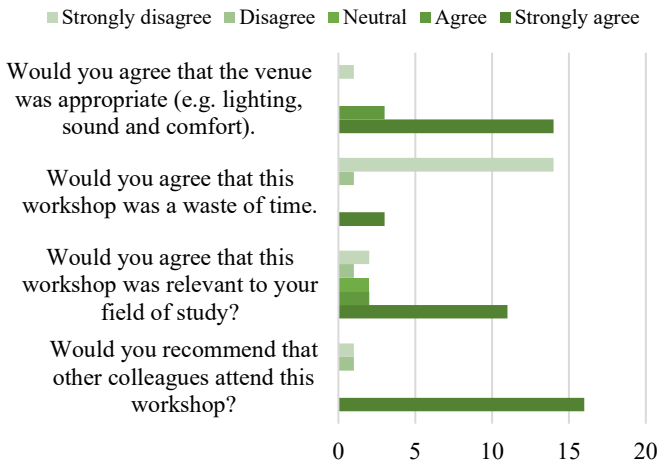


Figure 7: Participant perceptions regarding the workshop

Table 1: Questions used in the pre- and post-test

Number	Questions	Possible choices
Q5	Allow an AC signal to pass, while blocking a DC signal	Capacitor
Q6	Reduce the number of electrons flowing in a circuit	Resistor
Q7	Provide voltage regulation	Zener Diode
Q8	Block direct current in only one direction	Diode
Q9	Isolate two different circuits	Transformer
Q11	Measure the amount of water in a JOJO tank	Ultrasonic sensor
Q12	Determine if the door to your home is open	Reed Switch
Q13	Used as a smoke detector	Photocell
Q14	Switch on a light when a person enters a room	PIR receiver
Q15	Determine if a car's boot is open (no magnets available)	Tilt sensor

The first five questions (Q5 – Q9) focused on basic electronic components, such as the capacitor, resistor, Zener diode, diode and transformer. The next five questions (Q11 – Q15) focussed on sensor based electronic components, including the ultrasonic sensor, reed switch, photocell, PIR and tilt sensor. Participants were expected to engage with the majority of these components during their practical "hands-on" work in the laboratory. Knowledge of these components would be required before they can be correctly applied and connected to an Arduino microprocessor. The pre-test sought to establish the conceptual knowledge of the participants regarding these electronic components, which was then deconstructed (if their initial perceptions were wrong) and then re-constructed during the rest of the practical workshop. The post-test sought to establish whether the participant conceptual knowledge had been strengthened (if their initial perceptions were right) or had been re-constructed.

The pre-test was done at the start of the workshop using an ERS. Immediate feedback of all the participant responses was presented and analysed in order to strengthen or reconstruct relevant conceptual knowledge. The post-test was done at the end of the workshop using Google Forms, as the academics were in a computer-based laboratory with Internet access. The scores for the pre-test (dotted line) and post-test (solid line) for 2017 and 2018 are shown in Figure 8 and Figure 9 respectively.



Figure 8: Participant responses to the conceptual knowledge questions for 2017



Figure 9: Participant responses to the conceptual knowledge questions for 2018

Questions 5, 7 and 9 indicate that the workshop impacted positively on the conceptual knowledge of the participants regarding the capacitor (Q5), the Zener diode (Q7) and the transformer (Q9). Only a few participants could correctly select the transformer in the pre-test as a component which may be used to

isolate two different circuits. However, the majority of the participants selected the correct answer in the post-test. This suggests that the academics conceptual knowledge was reconstructed. No improvement was noticed regarding resistors (Q6) in both workshops (2017 and 2018). Participants also did not struggle in identifying the purpose of a diode (Q8) in 2017, although this did prove a little challenging in 2018 (3 out of 8 correctly identified the purpose with the pre-test). This may indicate that resistors and diodes are well known, and well used, by these participants.

Question 10 (Q10) was added in the second workshop to incorporate an inductor, as it was missing from the first workshop. Participants struggled to identify its purpose in the pre-test (4 out of 8) while seven participants correctly identified its purpose in the post-test for 2018.

Questions 11 through 15 relate to other electronic sensors that are commonly integrated with Arduino technology (see Table 1). Improvements in conceptual knowledge are evident for Q11 (Ultrasonic sensor), Q12 (Reed Switch) and Q14 (PIR). This reconstruction of knowledge can surely benefit these academics in terms of improving their own teaching practice. Participants strengthened their conceptual knowledge regarding the tilt sensor (Q15), as they correctly identified its purpose in the pre-test and in the post-test.

Pre-test results for the photocell in 2017 (Q13 in Figure 8) revealed that 8 out of the 10 academics would use it as a light sensor. The question was phrased as follows “Please select the component that you would use for a day/night sensor”. Obviously, the participants could associate the term “photo” with “light”, and thereby selected the photocell as the desired component. However, the post-test revealed that only 3 academics would use it as a smoke detector, which is a decrease of 50% from the pre-test with regard to the conceptual knowledge of the participants. The question in the post-test considered a different application for the photocell and was phrased as follows “Please select the component that you would use as a smoke detector”. The majority of the participants could not deduce that the presence of smoke would interrupt the reception of light by the photocell, thereby affecting its resistance and triggering an alarm condition. This indicated that more time needed to be allocated to the discussion of the photocell in future workshops and that the facilitators needed to revisit this question in the post-test.

This was done prior to the second workshop that was offered in 2018. The question was revised to state: “Please select the component that you would use to switch a light on when it becomes dark”. Phrasing the two questions differently, but using the same application, by using synonyms for the key words’ day (being light) and night (being dark) resonated more with the participants. Results in Fig. 9 show NO decrease in the participant’s conceptual knowledge for this component (Q13), but an actual strengthening of it. The correct response to the purpose of the photocell, or LDR, was observed for both tests in 2018. Using different applications (day/night sensor versus a smoke sensor) for the same component (photocell) proved a little problematic for the majority of the participants. This can further be addressed in the first stage of future workshops by discussing with the participants a greater variety of applications for each of the electronic components.

7. Conclusions

The purpose of this article was to illustrate how engineering academics from a TVET college were able to engage with the theory and practice of sensor integration with Arduino technology, thereby strengthening or reconstructing their conceptual knowledge of these electronic components. A practical workshop, or academic development workshop, was designed and facilitated by experts in electrical engineering. The 2-day workshop consisted of two stages that were presented to 18 academics from a TVET college in SA during 2017 and in 2018. The first stage sought to obtain the conceptual knowledge of TVET academics relating to 10 electronic components and sensors using a pre-test. The second stage related to four practical assignments that the participants needed to complete in a laboratory, followed by an online post-test that was again used to determine their conceptual knowledge. This was done to determine the value, and impact, of the practical workshop.

Conceptual knowledge regarding capacitors, Zener diodes, transformers, ultrasonic sensors, reed switches and PIR had to be deconstructed, and then reconstructed again during the workshop for many of the participants. However, their conceptual knowledge regarding photocells deteriorated in 2017, requiring the facilitators to review their discussion on it. This was done before the second workshop was offered in 2018. The results of the revised question showed no decrease in the participant’s conceptual knowledge for this component in 2018, but an actual strengthening of it. A key recommendation in this regard is to retain the same application for each electronic component between the pre-test and post-test. However, synonyms or different phrases could be used between the two questions. Another option is to enhance the discussion after the pre-test, but before the post-test, by providing participants with a greater variety of applications for each of the electronic components.

Conceptual knowledge regarding resistors, diodes and tilt sensors were strengthened, as the results between the pre- and post-test remained similar. TVET academics further indicated that the practical workshop was relevant to their work and that they would encourage fellow colleagues to attend future workshops.

A limitation of this study is related to the small sample size, where only 18 TVET college academics attended the workshops. However, the research methodology chosen was a case study that aims to explore and describe a setting where the researcher can gain a more in-depth understanding of what he/she is seeking. In this research, the facilitators were seeking to determine if the practical workshop (the setting) that they designed would be able to help TVET college academics to strengthen, or reconstruct, their conceptual knowledge of electronic components that may be integrated with the Arduino technology. This was achieved as only one question had to be rephrased in the pre and post-test. Subsequent offerings of this practical workshop should therefore not present any major challenges with regard to determining conceptual knowledge, as it has been validated on two occasions.

It is recommended that more such academic development workshops be offered as practical workshops by field experts to

enable fellow academics to either strengthen, or reconstruct, their conceptual knowledge of relevant technologies. This would also greatly contribute to meeting the requirements of CPD, as required by professional engineering bodies around the world.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors wish to acknowledge the financial support of the Manufacturing, Engineering and Related Services SETA (merSETA) in South Africa.

References

[1] Swart, A. J. and Hertzog, P. E., "Engaging academics from an engineering training college with Arduino sensors using an academic development workshop," presented at the ICST 2018, 12th International Conference for Sensing Technology, Limerick, Ireland, 2018.

[2] Grinter, L. E., "Responsibility in engineering education," The Journal of Higher Education, **25**(5), 258-261, 1954.

[3] Swart, A. J., "Does it matter which comes first in a curriculum for engineering students – theory or practice?," IJEEE, International Journal of Electrical Engineering Education, **47**(2), 189-199, 2010. DOI:10.7227/IJEEE.47.2.8.

[4] McCulloch, A. and Loeser, C., "Does research degree supervisor training work? The impact of a professional development induction workshop on supervision practice," Higher Education Research & Development, **35**(5), 968-982, 2016.

[5] Swart, A. J., Luwes, N., Olwagen, L., Greyling, C., and Korff, C., "Scholarship of teaching and learning – "What the hell" are we getting ourselves into?," EJEE, European Journal of Engineering Education, **42**(6), 653-667, 2016. DOI:10.1080/03043797.2016.1214689.

[6] Mardapi, D. and Herawan, T., "Community-Based Teacher Training: Transformation of Sustainable Teacher Empowerment Strategy in Indonesia," Journal of Teacher Education for Sustainability, **21**(1), 48-66, 2019. DOI:10.2478/jtes-2019-0004.

[7] Umelo, N., Amadi, A., Obodoeze, F., and Onyibe, C., "A Multi-purpose Hardware Efficient Temperature Regulator with LCD Display," International Journal of Computer Engineering and Information Technology, **9**(10), 254-257, 2017.

[8] Yüksekkaya, M. and Haberal, O. E., "An Efficient Microcontroller Course With An Affordable And Easy To Use Development Setup," presented at the EPESS, The Eurasia Proceedings of Educational & Social Sciences, 2015.

[9] Fan, Y. B. and Liu, H., "Study on Application Model of Internet of Things for Green Manufacturing," Applied Mechanics and Materials, **484**(1), 187-190, 2014. DOI:10.4028/www.scientific.net/amm.484-485.187.

[10] Vorster, J. and Quinn, L., "Privileging knowledge, creating knowers: An analysis of a formal programme for university lecturers," Re-Imagining Academic Staff Development: Spaces for Disruption, edited by L. Quinn), 71-88, 2012.

[11] Ndebele, C., "Deconstructing the narratives of educational developers on the enabling and constraining conditions in their growth, development and roles as educational staff development facilitators at a South African University," International Journal of Educational Sciences, **6**, 103-115, 2014. DOI:10.1080/09751122.2014.11890123.

[12] Makondo, L. and Makondo, O., "The Interplay between South African Higher Education Context and Academic Development," Anthropologist, **17**(2), 301-309, 2014. DOI:10.1080/09720073.2014.11891438.

[13] Swart, A. J., "Distance Learning Engineering Students Languish Under Project-Based Learning, But Thrive in Case Studies and Practical Workshops," IEEE Transactions on Education, **59**(2), 98-104, 2016. DOI:10.1109/TE.2015.2453349.

[14] Gibbs, G. R., "Workshop: Producing educational videos: design and workflow issues," presented at the Inspire Conference, HEA Social Sciences Conference, The Studio, Manchester, 2015.

[15] Ooi, P. C. and Tan, M. T., "Effectiveness of workshop to improve engineering students' awareness on engineering ethics," Procedia-Social Behavioral Sciences, **174**(2015), 2343-2348, 2015. DOI:10.1016/j.sbspro.2015.01.898.

[16] Manley, K., Martin, A., Jackson, C., and Wright, T., "A realist synthesis of effective continuing professional development (CPD): A case study of healthcare practitioners' CPD," Nurse education today, **69**, 134-141, 2018. DOI:10.1016/j.nedt.2018.07.010.

[17] Cole, M. J., "An Exploration of the Long-Term Impacts of Short CPD Workshops," PhD, Northumbria University, 2018.

[18] Mack, H. G., Golnik, K. C., Murray, N., and Filipe, H. P., "Models for implementing continuing professional development programs in low-resource countries," MedEdPublish, **6**, 1-18, 2017. DOI:10.15694/mep.2017.000018.

[19] Cruywagen, H., "Continuing professional development for the quantity surveying profession in South Africa:: commentary," Acta Structilia: Journal for the Physical Development Sciences, **14**(2), 91-103, 2007. DOI:hdl.handle.net/10520/EJC110013.

[20] Kasimba, R., "Training and Development Programs in Educational Institutions: A Human Factor Approach," Journal of Gleanings from Academic Outliers, **3**(1), 64-86, 2014.

[21] Altbach, P., "Higher education and the WTO: Globalization run amok," International Higher Education, **23**, 2-4, 2015.

[22] Shapiro, J. P. and Stefkovich, J. A., Ethical leadership and decision making in education: Applying theoretical perspectives to complex dilemmas. New York: Routledge. 2016.

[23] Spooner, E., Interactive Student Centered Learning: A Cooperative Approach to Learning. Lanham: Rowman & Littlefield. 2015.

[24] Ryan, M. and Ryan, M., "Theorising a model for teaching and assessing reflective learning in higher education," Higher Education Research & Development, **32**(2), 244-257, 2013. DOI:10.1080/07294360.2012.661704.

[25] Hertzog, P. E. and Swart, A. J., "Arduino - Enabling engineering students to obtain academic success in a design-based module! ," presented at the EDUCON 2016, IEEE Global Engineering Education Conference, Abu Dhabi, UAE, Dusit Thani Hotel, 2016.

[26] Swart, A. J. and Hertzog, P. E., "LED's as viable power loads for experimental purposes relating to PV modules in pico-solar systems," presented at the ICETAS 2017, AMA International University Bahrain, Bahrain, 2017.

[27] Swart, A. J. and Hertzog, P. E., "Evaluating the Performance of Small-Scale PV Modules in a Semi-Arid Area to Identify any Anomalies," IJST, Indian Journal of Science and Technology, **12**(28), 1-9, 2019. DOI:10.17485/ijst/2019/v12i28/147001.

[28] Love, P. E., Zhou, J., and Matthews, J., "Project controls for electrical, instrumentation and control systems: Enabling role of digital system information modelling," Automation in Construction, **103**, 202-212, 2019. DOI:10.1016/j.autcon.2019.03.010.

[29] Neafsey, P. J., "Computer-assisted instruction for home study: a new venture for continuing education programs in nursing," The Journal of Continuing Education in Nursing, **28**(4), 164-172, 1997. DOI:10.3928/0022-0124-19970701-06.

[30] Schlingloff, B.-H., "Towards a curriculum for model-based engineering of embedded systems," in Tagungsband des Dagstuhl-Workshops, H. Giese*, et al.*, Eds., ed Munchen: fortiss GmbH, 2014, pp. 83-90.

[31] Robertson, S.-A. and Graven, M., "Exploring South African mathematics teachers' experiences of learner migration," Intercultural Education, **26**(4), 278-295, 2015. DOI:10.1080/14675986.2015.1071754.

[32] Kirillov, A. V., Tanatova, D. K., Vinichenko, M. V., and Makushkin, S. A., "Theory and practice of time-management in education," Asian Social Science, **11**(19), 193-204, 2015. DOI:10.5539/ass.v11n19p193.

[33] Hori, T. and Alley, M., "W-10 Rethinking Oral Presentations for Engineers in a Global Society," presented at the JSEE, Japanese Society for Engineering Education Annual Conference, Tokyo City University, Setagaya Campus, 2017.

[34] Kolsaker, A., "Academic professionalism in the managerialist era: A study of English universities," Studies in Higher Education, **33**(5), 513-525, 2008. DOI:10.1080/03075070802372885.

[35] Hanssen, T.-E. S. and Solvoll, G., "The importance of university facilities for student satisfaction at a Norwegian University," Facilities, **33**(13), 744-759, 2015. DOI:10.1108/F-11-2014-0081.

[36] Buthelezi, Z. and Training, "Lecturer experiences of TVET College challenges in the post-apartheid era: a case of unintended consequences of educational reform in South Africa," Journal of Vocational Education, **70**(3), 364-383, 2018. DOI:10.1080/13636820.2018.1437062.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 460-468 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

The Usability Evaluation of Academic Progress Information System (SIsKA-NG)

Gede Indrawan*, I Made Agus Oka Gunawan, Sariyasa

Computer Science Department, Universitas Pendidikan Ganesha, 81116, Indonesia

ARTICLE INFO

Article history:
Received: 27 January, 2020
Accepted: 22 February, 2020
Online: 04 April, 2020

Keywords:
Performance Measurement
Retrospective Think Aloud
User Experience Questionnaire

ABSTRACT

Limited functionalities of the Academic Progress Information System (SIsKA) has direct consequences on the quality of thesis management service at the authors' magister study program of the authors' university. This research focused on the significant improvement from the previous User Experience Questionnaire (UEQ) result of SIsKA. That significant improvement was based on the recommendations from the previous Heuristic Evaluation (HE) and two additional usability evaluation methods. All of those improvements have been applied to the new Academic Progress Information System that was called SIsKA Next Generation (SIsKA-NG). Overall, there are three usability evaluation methods were used in this research, namely the Performance Measurement (PM), the Retrospective Think Aloud (RTA), and the UEQ. The PM through the effectiveness analysis and the efficiency analysis each provided 11 and 16 recommendations, while The RTA provided 39 recommendations. The UEQ provided the assessment on SIsKA-NG with level excellent on aspect of attractiveness, efficiency, dependability, stimulation, and novelty; and with level good on aspect of perspicuity. It was considered as a significant improvement of the UEQ result of SIsKA-NG compared to the previous UEQ result of SIsKA because of the better assessment on four aspects with level excellent on SIsKA-NG, namely attractiveness, efficiency, dependability, and novelty. Same assessments are on two aspects, namely perspicuity and stimulation. SIsKA-NG still need to be improved based on the aspect of perspicuity, additional recommendations from the conducted Focus Group Discussion, and procedural White Box testing for code evaluation.

1. Introduction

This paper conformed to the scope of the journal in Computer software and applications (A6) of Engineering & Technology topic. It addressed a particular complex system related to the thesis management information system. This paper is an extension of the work originally presented as a best paper in The 3rd Third International Conference on Informatics and Computing (ICIC) [1]. This paper focused on the significant improvement from the previous User Experience Questionnaire (UEQ) result of SIsKA. That significant improvement was based on the recommendations from the previous Heuristic Evaluation (HE) [1] and two additional usability evaluation methods conducted in this research, namely the Performance Measurement (PM) and the Retrospective Think Aloud (RTA). All of those improvements have been applied to the new Academic Progress Information System that was called SIsKA Next Generation (SIsKA-NG). SIsKA itself is Indonesian abbreviation of "Sistem Informasi Kemajuan Akademik".

SIsKA that has been used at the authors' magister study program (refer to " the authors' study program " on the next discussion) of the authors' university has been able at certain level to manage the academic progress of the postgraduate student (refer to "the student" on the next discussion) related to their thesis research [1][2]. The future works of that previous research are on improving the UEQ result and following up the user interface improvement recommendations based on the HE using expert respondents.

This research focused on that future works and following up the recommendations from two additional usability evaluation methods, namely the PM and the RTA. The usability evaluation by involving the user respondents has an advantage on information about how the users use the system as well as problems faced directly by them [3]. There are three categories of usability

* Gede Indrawan, 685 Email : gindrawan@undiksha.ac.id

evaluation methods, namely inspection, testing, and inquiry [4]. This research used both of testing and inquiry for observing the users and their response to get the improvement recommendations.

This paper is organized into several sections, i.e.: Section 1 (Introduction) describes the problem background related to the improvement of SIsKA based on its previous UEQ result [1]; Section 2 (Literature Survey) describes the related works in area of usability testing methods; Section 3 (Research Method) contains the research stages for all of SIsKA improvements applied to SIsKA-NG; Section 4 (SIsKA-NG) provides several aspects related to its initial design and implementation; Section 5 (Result and Discussion) covers SIsKA-NG usability evaluation; and finally, Section 6 (Conclusion and Future Work) consists of some important concluded points.

2. Literature Survey

There are many usability testing methods that have been developed and tested to provide recommendations for improving a system. In one study, Ita et al. [1] successfully utilized the HE to produce recommendations based on expert aspects in the interface field. The PM is another usability testing method, where in one research, Widyanti et al. [5], used it in evaluating the interface of online transportation. The PM was used to obtain quantitative data to analyze the level of effectiveness and efficiency when the users carry out certain tasks [4][5][6]. Widyanti et al. combined the PM with the RTA, where the RTA is a Think Aloud method that can be combined with other usability methods [7][8]. Elling et al. [8] combined with eye movements method in analyzing the behavior of system users. Those research [5][8] provided a list of recommendations for the improvement of the system being evaluated.

In addition to usability testing, the level of satisfaction of user experience also needs to be calculated by determining the aspects to be evaluated [1]. The inquiry method is one method that can be used to evaluate user experience [4]. The UEQ, as one type of Questionnaire, usually takes 3-5 minutes to read and complete the questionnaire [9]. The UEQ has the questionnaire items that can make the respondents do not think too long to fill out the questionnaire [10]. The purpose of the UEQ is to compare the level of user experience between two products, test the user experience of a product, and determine area of improvement [11]. The UEQ has the advantage of measuring aspects of user experience of the product very quickly [12].

In our research, we combined the PM method with the RTA method by involving system users. We also calculated the satisfaction value of user experience using UEQ to get aspects of user experience that must be improved. This research is not limited to obtain system improvement recommendations as in the previous studies [1][5][8], but also carried out the implementation phase at the authors' postgraduate program server. So that the recommendations for improvements that are obtained directly can be tested by the users and of course will get other recommendations, closely related to the system implementation environment.

3. Research Method

Figure 1 shows the research method in the form of a flow chart. The initial stage of this research was to design and to implement SIsKA-NG that follows up the previous HE recommendations [1] related to the user interface improvement of SIsKA web page (refer to "page" on the next discussion), as shown by Table 1. Note that Table 1 has been updated by additional expert recommendations since the previous publication [1].

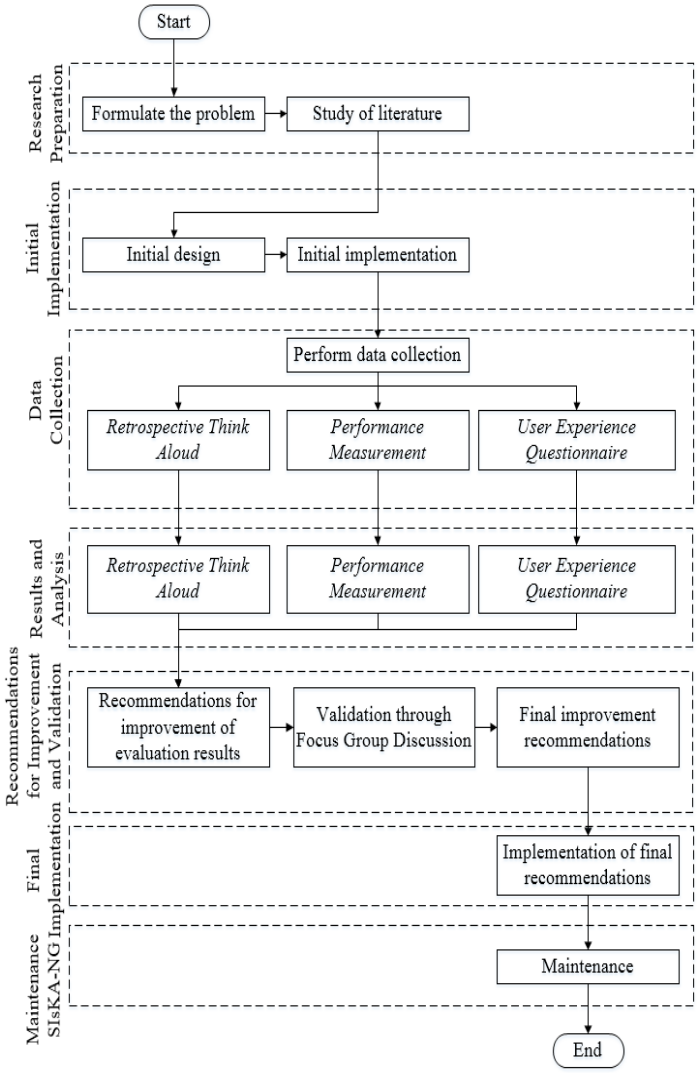


Figure 1: Research Method

The usability evaluation with the PM and the RTA was conducted afterwards, and their improvement recommendations were then validated by SIsKA-NG manager, namely the authors' study program manager and the authors' postgraduate program Information Technology (IT) manager. The process was continued to implement the recommendations on SIsKA-NG at the authors' postgraduate program server. The maintenance was conducted after establishing the implementation.

The UEQ used 20 respondents [11][13] who carried out certain tasks according to the pages being evaluated [14]. Evaluation on the administrator pages used 25 tasks, the lecturer pages used 14 tasks, and the student pages used 21 tasks. Each respondent run each task 2 times to get a difference in the level of effectiveness and efficiency. The final stage of data collection was filling in the questionnaire (Figure 2) combined with the additional questions to clarify the answers of the respondents [15].

Table 1: The Updated Previous Recommendations of SIsKA

No	Violated principle(s)	Recommendation(s)
1	Visibility of system status (feedback). When the confirmation to delete the data appears, the "Delete" button is blue, while the "Back" button is red.	1. Use standard color of buttons, such as red for "Delete" button and blue for "Back", so the users do not misunderstand and make a mistake. 2. There should be a response that distinguishes visually when the object is given an action (selected, pressed, etc.). 3. Naming the menu and the page in accordance with its content. 4. Color must be clear on the button or functional system.
2	Use Control and Freedom. The users are forced to change the password if the password is the same with the username.	1. Instead of forcing the users to change the password, they should be given a warning. The users have the freedom to decide either to change or to leave their password. 2. There should be a feature to manage personal account data of each user, including sensitive information such as password.
3	Flexibility and Efficient of Use. Search feature does not exist yet, there is only a search in the table, which is not appropriate.	The search feature should be placed at the upper right corner of the page according to the standard system interface in general.
4	Recognation rather than Recall. There is no error warning message when the users make a mistake.	1. The system do not prevent the users from making a mistake. For an example, in making examination schedule. If the users have to add schedule first, the button to manage the participants should be disabled so that they are prevented to press the button. 2. The added schedule feature is done one by one, so that the users do not make a mistake, such as pressing the manage button of the participant first and then creating the examination schedule data.
5	Consistency and Standards. The "Add" menu is inconsistent.	1. The menu should be consistent, should add "Add" menu in lectures data page. 2. Consistency on the use of grammar in naming menus provided by the system. 3. Consistency on color selection on the system.
6	Aesthetic and Minimalist Design. There is no clear instructions/description on any given page, so the users can only guess the content/ intent of the page.	Each page is given a clear and informative description so that the users do not guess the intent of the page. For an example, the student home page still empty. It should be filled with the informative information, such as the registration flow of the proposal to the thesis examination; or the information related to the history of the stage passed by the student (see next Figure 4).
7	Error Prevention. The system do not prevent the users from making a mistake.	1. The system needs to prevent the users from making a mistake, such as providing a message to make sure whether they actually logout or not. 2. Provision of automatic data entry feature or the system provides clear data format, such as date, time, and telephone number. 3. Provision of warning feature about data or action that can cause error.
8	Help and Documentation. There is no help menu.	1. Creating a help menu to make it easier for the users to use the system, including: 1) News; 2) System Guidance; and 3) Frequently Asked Question (FAQ). 2. Creating a site map to make it easier for the users to navigate the system.
9	Match between system and real world	The use of grammar that can be understood directly by the users in accordance with their expectations.

		1	2	3	4	5	6	7		
The product should look enjoyable	annoying	○	○	○	○	○	○	○	enjoyable	1
The product should be understandable	not understandable	○	○	○	○	○	○	○	understandable	2
The product should be creative	creative	○	○	○	○	○	○	○	dull	3
The product should be easy to learn	easy to learn	○	○	○	○	○	○	○	difficult to learn	4
Using the product should be valuable	valuable	○	○	○	○	○	○	○	inferior	5
Using the product should be exciting	boring	○	○	○	○	○	○	○	exciting	6
Using the product should be interesting	not interesting	○	○	○	○	○	○	○	interesting	7
The interaction with the product should be predictable	unpredictable	○	○	○	○	○	○	○	predictable	8
I should perform my tasks with the product fast	fast	○	○	○	○	○	○	○	slow	9
The product should be inventive	inventive	○	○	○	○	○	○	○	conventional	10
The interaction with the product should be supportive	obstructive	○	○	○	○	○	○	○	supportive	11
The product should look good	good	○	○	○	○	○	○	○	bad	12
The product should be easy	complicated	○	○	○	○	○	○	○	easy	13
The product should look pleasing	unlikable	○	○	○	○	○	○	○	pleasing	14
The product should be leading edge	usual	○	○	○	○	○	○	○	leading edge	15
The product should look pleasant	unpleasant	○	○	○	○	○	○	○	pleasant	16
The interaction with the product should be secure	secure	○	○	○	○	○	○	○	not secure	17
Using the product should be motivating	motivating	○	○	○	○	○	○	○	demotivating	18
The interaction with the product should be meet my	meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
I should perform my tasks with the product efficient	inefficient	○	○	○	○	○	○	○	efficient	20
The product should be clear	clear	○	○	○	○	○	○	○	confusing	21
I should perform my tasks with the product practical	impractical	○	○	○	○	○	○	○	practical	22
I should perform my tasks with the product organized	organized	○	○	○	○	○	○	○	cluttered	23
The product should look attractive	attractive	○	○	○	○	○	○	○	unattractive	24
The product should look	friendly	○	○	○	○	○	○	○	unfriendly	25
The product should be innovative	conservative	○	○	○	○	○	○	○	innovative	26

Figure 2: Twenty-Six Statement Items of UEQ

4. SIsKA-NG

Based on Table 1, the initial design of SIsKA-NG with its main functionalities is elaborated by using a use case diagram [16], as shown by Figure 3. The use case diagram shows the main functionalities related to the management of research, involving three stages of examination (proposal, pre-thesis, and thesis), and three kinds of user (administrator, lecturer, and student). Part of those examination stages was shown by Figure 4. Related to kinds of user, the administrator manage any data related to the student research, including research data, examination prerequisites, and examination schedule. The students can submit their research data and examination prerequisites, and later view examination schedule, as well as list of existing researchs. The lecturers have access to see the examination schedule related to their role as a supervisor or an examiner. All those described functionalities can be done by the users if they have logged in to SIsKA-NG.

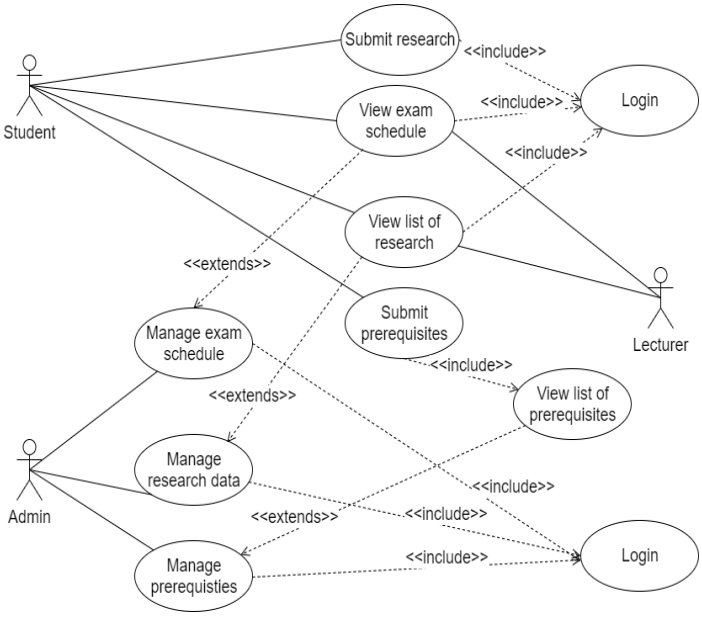


Figure 3 Use Case Diagram of SIsKA-NG

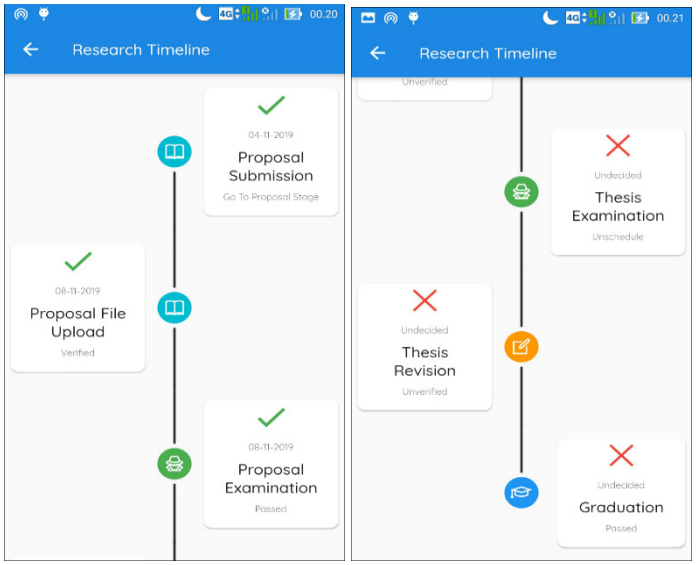


Figure 4 Part of the Student's Examination Stages of SIsKA-NG on its mobile application: Proposal (Left); Thesis (Right)

Figure 5 shows an the Entity Relationship Diagram (ERD) of SIsKA-NG with the entities and their relationships that are described by using the Crow's Foot model [17] from the previous recommendations [1], as shown by Table 1. There are 17 entities related to each other. These entities in the ERD is used for SIsKA-NG new database implementation.

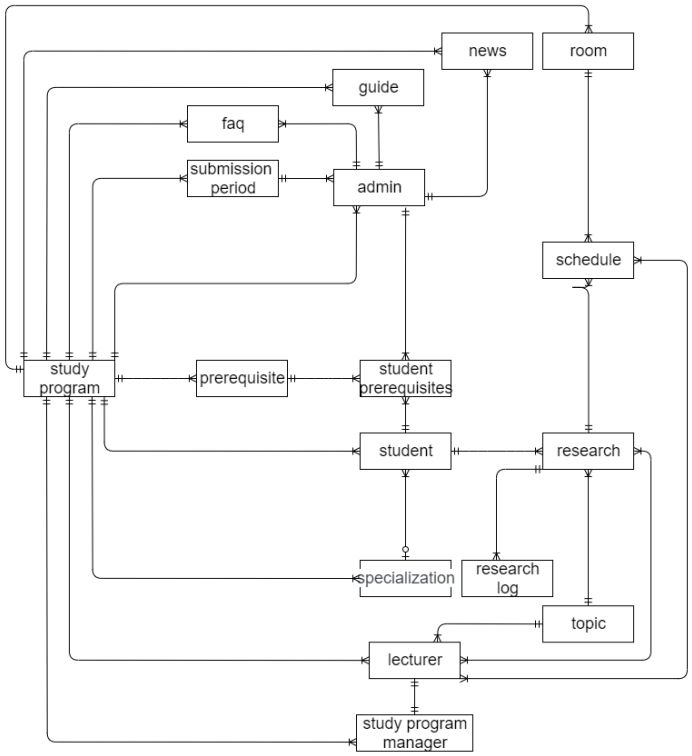


Figure 5 ERD of of SIsKA-NG

The ERD also shows relationship type of one-to-one and one-to-many. For example, one-to-one relationship connects the lecturer and the study program manager since only one lecturer can be a study program manager, meanwhile one-to-many relationship

connects the specialization and the student since one specialization can be assigned to many students.

In addition to the functionality requirement design using use case diagram and ERD, an analysis was also carried out for non-functionality requirement related to the implementation. Table 2 shows the non-functionality requirement that are tailored to the authors' postgraduate program server specification. Based on Table 2, a web framework CodeIgniter was used to improve the novelty aspect of SIsKA-NG technology. That framework supports PHP 5.6 and one of the latest web technologies [18][19].

Table 2 Non-Functionality Requirement of SIsKA-NG

No	Requirement	Specification
1	Web Server	Apache: 2.4.x
2	DBMS	MySQL: 5.x
3	PHP	PHP: 5.6

The result of the initial implementation based on the previous recommendations (Table 1) produced SIsKA-NG with an improved interface. As an example, Figure 6 shows the administrator starting page.



Figure 6: SIsKA-NG initial implementation of the Administrator Starting Page

5. Result and Discussion

5.1. The Performance Measurement

The PM was conducted through the effectiveness analysis and the efficiency analysis. The effectiveness of SIsKA-NG was obtained by counting the number of failures that occur when the respondent completes the task. The effectiveness analysis was conducted by reviewing the video of the work done by the respondent to see the factors that cause the respondent failed to complete the task. All factors that cause failure were then summarized so get a final list of causes of failure and recommendations for the improvement.

Table 3 shows the result of SIsKA-NG effectiveness analysis. On the administrator pages, there are four factors that cause failure, namely: 1) Unclear menus that need to be regrouped and renamed; 2) Less effective position of the student prerequisites sub menu; 3) Less effective position of the action button; and 4) Less effective display button. On the lecturer page, there are four factors that cause failure, namely: 1) Less effective position of the action button; 2) Less clear icon of the action button; 3) Less clear filter feature of each research stage and the examination schedule; and 4) The clarity of the grammar on the page. On the student page, there are four factors that cause failure, namely: 1) Less effective position of the action button; 2) Less clear icon of the action button; 3) The examination schedule that does not directly displays the detail schedule; and 4) The clarity of the grammar on the page. Table 4 shows the list of the recommendations based on the effectiveness analysis.

Table 3: The Result of SIsKA-NG Effectiveness Analysis

Pages	1st Test		2nd Test	
	Errors	Percentage	Errors	Percentage
Administrator	23	4,60%	5	1%
Lecturer	20	7,14%	8	2,86
Student	14	3,33%	9	2,14%

The efficiency of SIsKA-NG was obtained by comparing the time spent by the respondent on the first and the second trial. This efficiency calculation used a statistical comparison of the Mann Whitney U-test. Comparison of work time data from two experimental groups requires two hypotheses for each task, namely:

- H0 : There is no time difference in the completion of task **[task number]** between the first trial and the second trial.
- H1 : There is time difference in the completion of task **[task number]** between the first trial and the second trial.

Final computation of the efficiency data processing was done by comparing the *p*-value generated from the Mann Whitney U-test [6] of each task with a value of α equal to 0.05. If the *p*-value is greater than 0.05, the decision taken do not reject H0. Analysis of the factors causing the efficiency is also done if H1 is true. Analysis of the factors causing this level of efficiency is done by reviewing the video of the work done by the respondent to see the factors causing the difference in task completion time. All factors are then summarized to obtain a final list of causes of the time difference in efficiency and its recommendations for the improvement.

Based on the results of *p*-value comparison of each task with value α, then statistically there are 4 tasks that do not have a significant difference and 21 other tasks that have a significant difference from the completion time of each task on the administrator page. On the lecturer page, there are 10 tasks that do not have a significant difference and 4 other tasks that have a significant difference from the completion time of each task. On the student page, there are 10 tasks that do not have a significant difference and 11 other tasks that have a significant difference from the completion time of each task.

Tasks that show a significant difference in processing time are then analyzed in terms of causal factors by re-observing the video of respondent's work. It was found that on the administrator page, many respondents have difficulty in finding the action buttons and menus grouping that are unclear, so that in the first trial more time was needed to search for menus or functions for task completion.

On the lecturer and the student pages, significant time differences in some assignments were also found caused by inconsistent and unclear buttons positions, besided unclear grammar. Table 5 shows the list of the recommendations based on the efficiency analysis.

Table 4: The Recommendations based on the Effectiveness Analysis

No	Pages	The Recommendations
1	Administrator	1. Be clear on the group and the name of the menus related to their function. 2. The student prerequisites sub menu should be moved to the main menu. 3. The action button should be moved to the left position of the page. 4. The display button should be removed so the users only need to input the filter data.
2	Lecturer	1. The action button should be moved to the left position of the page. 2. The icon of the action button should be easy to understand. 3. Be clear on filter feature at each research stage and exam schedule. 4. Be clear on the grammar on the page.
3	Student	1. The action button should be moved to the left position of the page. 2. The icon of the action button should be easy to understand. 3. The examination schedule should directly displays the detail schedule. 4. Be clear on the grammar on the page.

Table 5: The Recommendations based on the Efficiency Analysis

No	Pages	The Recommendations
1	Administrator	1. Be clear on the group and the name of the menus related to their function. 2. The action button should be moved to the left position of the page. 3. The save button should be moved to the bottom position of the page. 4. The display button should be removed so the users only need to input the filter data. 5. The entire examination schedule should be displayed and could be filtered afterwards. 6. Active Menu Tab should be colored for the visual clarity to the users. 7. Help text should be added for the usage clarity to the users.
2	Lecturer	1. The save button should be moved to the bottom position of the page. 2. Filter feature should be added to the stages in all research submenus. 3. The display button should be removed so the users only need to input the filter data.
3	Student	1. Be clear on the group and the name of the menus related to their function. 2. The action button should be moved to the left position of the page. 3. The save button should be moved to the bottom position of the page. 4. Summary of the previous research stage is automatically displayed when the users add research data. 5. The examination schedule should directly displays the detail schedule. 6. Active Menu Tab should be colored for the visual clarity to the users.

5.2. The Retrospective Think Aloud

Table 6, Table 7, and Table 8 show the summary of difficulties, problems, suggestions, and/or criticisms on SIsKA-NG page of the administrator, the lecturer, and the student, respectively. Related to the position mentioned further, it refered to the position at SIsKA-NG page.

Table 6: Summary of the RTA on the Administrator Pages of SIsKA-NG

No	Respondent Code	Summary
1	R01, R04, R05, R06, R08, R10, R12, R14, R15, R16, R17, R19, R20	Features/aspects on the main menu and perspicuity: 1. Menus seem unclear and should be grouped based on their similar function. 2. Menus should stay visible so that when the users are at the bottom position, there is no need to scroll up to find them. 3. Information of each form's field should be added to avoid user confusion. 4. Alternating background color of row should be used for the visual clarity of data row. 5. Tabs at the revision section of each research stage seem unclear and should be filled with distinguishing colors. 6. Error notifications will stay visible, with the users having the freedom to close them.
2	R01, R02, R03, R05, R06, R07, R08, R09, R11, R13, R14, R15, R17, R18, R19, R20	Features/aspects of the system button: 1. The save button should be moved to the bottom position of the page. 2. The system functional button should be moved to the left position of the page. 3. The display button should be removed for automatic data display.
3	R01, R02	Features/aspects on the information menu: 1. Content should be enriched. 2. The FAQs should be categorized specifically for the lecturer and the student.
4	R08, R09	Features/aspects on the account setting:

No	Respondent Code	Summary
		1. Account feature should be placed at the upper-right position for the easy access. 2. Logged-in user display should be clickable to access the account management page.
5	R05	Feature/aspect on the student prerequisites: its menu should be grouped according to each research stage.
6	R08	Feature/aspect on the file upload: there should be a preview when uploading image data.
7	R11	Feature/aspect on the system bug: data sorting fixation on the submission page.

Table 7: Summary of the RTA on the Lecturer Page of SIsKA-NG

No	Respondent Code	Summary
1	R01, R04, R05, R06, R07, R12, R13, R17, R18, R20	Features/aspects on the data filter: 1. Filter feature should be added to all menus at the research stages. 2. Filter feature should be added according to the examination role.
2	R02, R05, R06, R08, R09, R10, R12, R14, R15, R17, R18, R19, R20	Features/aspects of the system button: 1. The icon of the button should be easy to understand. 2. The action button should be moved to the left position of the page. 3. The display button should be removed on the schedule page.
3	R03	Feature/aspect on the information menu: content should be enriched.
4	R03, R08	Feature/aspect on the account setting: account feature should be placed at the upper-right position for the easy access.
5	R11, R14, R16	Feature/aspect on the main menu: There should be a separation between menu groups for the visual clarity to the users.

Table 8: Summary of the RTA on the Student Page of SIsKA-NG

No	Respondent Code	Summary
1	R01, R03, R05, R06, R07, R08, R10, R11, R12, R14, R15, R16, R17, R19	Features/aspects on the main menu and perspicuity: 1. Menus seem unclear and should be grouped based on their similar function. 2. Menus should stay visible so that when the users are at the bottom position, there is no need to scroll up to find them. 3. Information of each form's field should be added to avoid user confusion. 4. Unnecessary fields should be removed to avoid user confusion. 5. Tabs at the revision section of each research stage seem unclear and should be filled with distinguishing color. 6. Error notifications will stay visible, with the users having the freedom to close them. 7. Each stage in the timeline should be given a distinguishing color for the visual clarity of the stages that have been passed (see Figure 4).
2	R01, R02, R03, R04, R05, R06, R07, R08, R09, R11, R13, R14, R15, R18, R19, R20	Features/aspects of the system button: 1. The save button should be moved to the bottom position of the add or edit page. 2. The action button should be moved to the left position of the start page of each menu.
3	R02	Feature/aspect on the information menu: content should be enriched.
4	R08, R09	Features/aspects on the account setting: 1. Account feature should be placed at the upper-right position for the easy access. 2. Logged-in user display should be clickable to access the account management page.
5	R10, R13, R18	Features/aspects on the research stage: summary from the previous stage is automatically seen when adding research data of the recent stage.
6	R07, R17	Features/aspects on the student prerequisites: 1. The student only need to upload prerequisite files according to the prerequisite requirement. 2. Its menu should be grouped according to each research stage.
7	R04, R11, R12, R18, R20	Feature/aspect on the examination schedule: detail schedule for the logged-in student should be displayed directly.

5.3. The User Experience Questionnaire

After following up the PM and the RTA recommendations, SIsKA-NG was used for the UEQ. The questionnaire was filled out by 20 respondents using SIsKA-NG. The UEQ value of 26 statement items (see Figure 2) obtained from each respondent was processed by using the provided UEQ Data Analysis Tool [9][13]. At the data analysis stage, testing was carried out by comparing the value of each aspect with the product data set available in the UEQ Analysis Data Tool. Benchmark test can describe the relative quality of SisKA-NG compared to the other products. Benchmark test result are divided into five categories, namely *Excellent*, *Good*, *Above Average*, *Below Average*, and *Bad*. Data obtained from the UEQ was the result of SIsKA-NG user experience measured

through six aspects, namely *attractiveness*, *perspicuity*, *efficiency*, *dependability*, *stimulation*, and *novelty* [11].

Figure 7 shows the result where Good category was obtained for *perspicuity* aspect. This category means that 25% of the products in the dataset are better than SIsKA-NG while 50% of the others are worse. Excellent category was obtained for aspect of *attractiveness*, *efficiency*, *dependability*, and *stimulation*. Those results are in the range of 10% best results. Excellent category was also obtained for novelty aspect but 10% of the products in the dataset are better than SIsKA-NG while 75% of the others are worse.

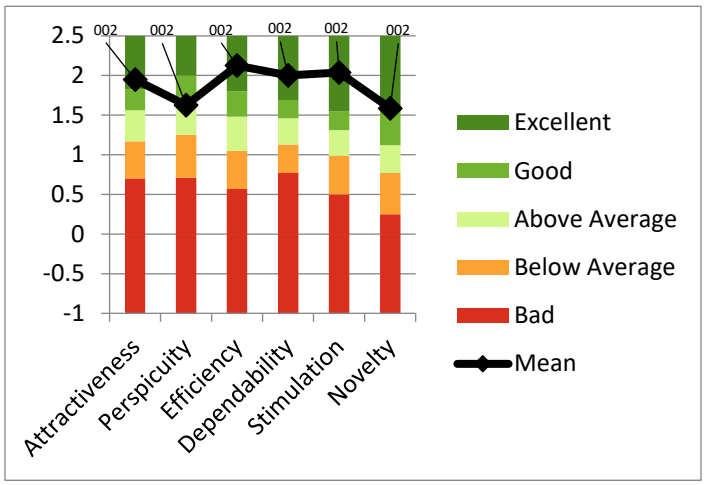


Figure 7: SIsKA-NG Benchmark Graph

Previous SIsKA with its six aspects each was assessed with the UEQ by score 1.59; 1.75; 1.64; 1.40; 1.56; and 1.10 [1]. The UEQ result of SIsKA-NG, as shown by Figure 7, was considered as a significant improvement compared to the previous UEQ result of SIsKA because of the better assessment on four aspects with level excellent, namely *attractiveness*, *efficiency*, *dependability*, and *novelty*. Same assessments are on two aspects, namely *perspicuity* (at level good eventhough SIsKA has higher score) and *stimulation*.

5.4. The Focus Group Discussion

Table 9 shows the additional recommendations that come from the Focus Group Discussion (FGD).

Table 9: The additional recommendations based on the FGD

No	The Recommendations
1	Addition of CSV file import feature for user data.
2	Printing of the examination schedule should be more attractive and based on existing design.
3	Addition of verification feature by the research supervisor.
4	Integration of the student prerequisites with the research submission.

The FGD discussed the result of the improvement recommendations obtained from the PM and the RTA evaluation. That FGD was conducted with the SIsKA-NG administrator, namely the authors' study program administrator and the authors' postgraduate program IT administrator. Those additional recommendations at Table 9 were considered as a reference for the further implementation.

5.5. The Implementation

Figure 8 shows an example of the new interface of SIsKA-NG, as one of the results of the final implementation based on the recommendations related to the menus grouping and the menus naming. That new interface is the result of the page improvement at Figure 6, which is the administrator starting page. Based on point 1 of the effectiveness recommendations of the administrator pages (Table 4), point 1 of the efficiency recommendations of the administrator pages (Table 5), and the RTA result, an improvement was made by grouping and by naming the menus related to their function.

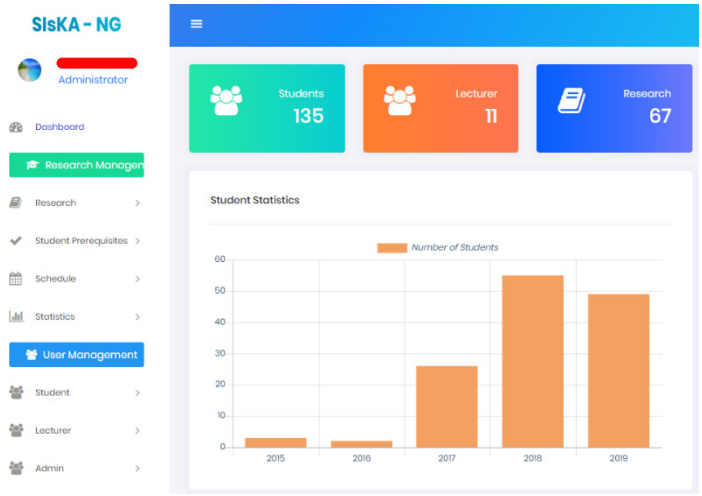


Figure 8 SIsKA-NG final implementation of the Administrator Starting Page

5.6. The Maintenance

SIsKA-NG, which has been hosted at the authors' postgraduate program server, has been successfully used for the real case of student research data management. However, there are several things to be noted during the use, including:

- The framework code used was unstable at the authors' postgraduate program server in the code of Asynchronous JavaScript And XML (AJAX) [20][21] for displaying the initial table (Figure 9). This error occurred because some Internet Servive Providers (ISPs) injected the advertising code that cause the AJAX code did not run perfectly. Based on in-depth testing, to avoid insertion of those advertisements SIsKA-NG need to use the Hypertext Transfer Protocol Secure (HTTPS) Internet communication protocol.
- The authors' postgraduate program server need to be configured since the development Operating System (OS) was not the same as the production OS, neither do some software libraries needed (Figure 10).
- The file size that can be uploaded need to be adjusted because of recent 4MB limitation.

- Email notifications did not run at the server side, so additional email setting is required at the authors' postgraduate program server.
- The time to run a feature of SIsKA-NG need to be optimized by evaluating the structure of SIsKA-NG code.

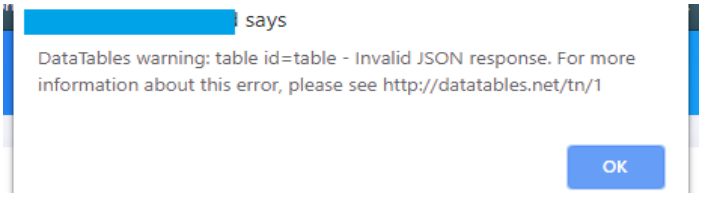


Figure 9: AJAX Error of SIsKA-NG

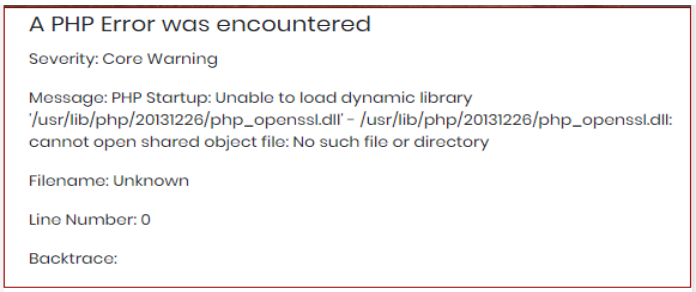


Figure 10: Software Library Error of SIsKA-NG

6. Conclusion and Future Work

This research has conducted significant improvement on the Academic Progress Information System that was called SIsKA-NG based on the usability evaluation. It gave better UEQ result of SIsKA-NG compared to the previous UEQ result of SIsKA since there are better excellent level assessment on four aspects of SIsKA-NG, namely *attractiveness*, *efficiency*, *dependability*, and *novelty*. Same assessments are on two aspects, namely *perspicuity* and *stimulation*. SIsKA-NG still need to be improved based on the aspect of *perspicuity*, additional recommendations from the conducted FGD, and procedural White Box testing for code evaluation.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgment

The authors would like to thank to Indonesian Ministry of Education and Culture that supported this work through research grant.

References

[1] A. I. I. Paramitha, G. R. Dantes and G. Indrawan, "The Evaluation of Web Based Academic Progress Information System Using Heuristic Evaluation and User Experience Questionnaire (UEQ)," in *Third International Conference on Informatics and Computing (ICIC)*, 2018. https://doi.org/10.1109/IAC.2018.8780430.

[2] G. Indrawan, G. T. Heriawan, A. A. I. I. Paramitha, G. Wiryawan, G. B. Subawa, M. T. Sastradi, and K. A. Sucahyana, "SIsKA: Mobile Based Academic Progress Information System," in *International Conference on Innovative Research Across Disciplines (ICIRAD)*, 2017. https://doi.org/10.2991/icirad-17.2017.24.

[3] J. M. C. Bastien, "Usability testing: a review of some methodological and technical aspects of the method" *Int. J. Med. Inform.*, **79**(4), e18–e23, 2010. https://doi.org/10.1016/j.ijmedinf.2008.12.004

[4] S. Gupta, "A Comparative study of Usability Evaluation Methods" *Int. J. Comput. Trends Technol.*, **22**(3), 103–106, 2015. https://doi.org/10.14445/22312803/IJCTT-V22P121

[5] A. Widyanti and S. A. Q. Ainizzamani, "Usability evaluation of online transportation' user interface" in *International Conference on Information Technology Systems and Innovation (ICITSI)*, 2017. https://doi.org/10.1109/ICITSI.2017.8331762

[6] F. Z. Ghazizadeh and S. Vafadar, "A quantitative evaluation of usability in mobile applications: An empirical study" in *International Symposium on Computer Science and Software Engineering Conference (CSSE)*, 2017. https://doi.org/10.1109/CSICSSE.2017.8320120

[7] S. Elling, L. Lentz, and M. de Jong, "Retrospective think-aloud method: Using eye movements as an extra cue for participants' verbalizations" in *Proceedings of the 2011 annual conference on Human factors in computing systems*, 2011. https://doi.org/10.1145/1978942.1979116

[8] M. W. M. Jaspers, "A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence" *Int. J. Med. Inform.*, **78**(5), 340–353, 2009. https://doi.org/10.1016/j.ijmedinf.2008.10.002

[9] H. Santoso, M. Schrepp, R. Y. Kartono Isal, A. Yudha Utomo, and B. Priyogi, "Measuring the User Experience" *J. Educ. Online*, **13**(1), 58–79, 2016. https://doi.org/10.9743/JEO.2016.1.5

[10] B. Laugwitz, T. Held, and M. Schrepp, "Construction and Evaluation of a User Experience Questionnaire" in *Holzinger A. (eds) HCI and Usability for Education and Work. USAB 2008. Lecture Notes in Computer Science, vol 5298. Springer, Berlin, Heidelberg*, 2008. https://doi.org/10.1007/978-3-540-89350-9_6

[11] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Construction of a Benchmark for the User Experience Questionnaire (UEQ)" *Int. J. Interact. Multimed. Artif. Intell.*, 4(4), 40-44, 2017. https://doi.org/10.9781/ijimai.2017.445

[12] D. S. S. Sahid, P. I. Santosa, R. Ferdiana, and E. N. Lukito, "Evaluation and measurement of Learning Management System based on user experience" in *The 6th International Annual Engineering Seminar (InAES)*, 2016. https://doi.org/10.1109/INAES.2016.7821910

[13] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios" in *In: Marcus A. (eds) Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. DUXU 2014. Lecture Notes in Computer Science, vol 8517. Springer, Cham*, 2014. https://doi.org/10.1007/978-3-319-07668-3_37

[14] J. Nielsen, T. Clemmensen, and C. Yssing, "Getting access to what goes on in people's heads?" in *Proceedings of the second Nordic conference on Human-computer interaction*, 2002. https://doi.org/10.1145/572020.572033

[15] A. Hinderks, M. Schrepp, F. J. Domínguez Mayo, M. J. Escalona, and J. Thomaschewski, "Developing a UX KPI based on the user experience questionnaire" *Comput. Stand. Interfaces*, **65**, 38–44, 2019. https://doi.org/10.1016/j.csi.2019.01.007

[16] M. Seidl, M. Scholz, C. Huemer, and G. Kappel, "The Use Case Diagram" in *UML @ Classroom. Undergraduate Topics in Computer Science. Springer, Cham*, 2015. https://doi.org/10.1007/978-3-319-12742-2_3

[17] S. Hitchman, "The Details of Conceptual Modelling Notations are Important - A Comparison of Relationship Normative Language" *Commun. Assoc. Inf. Syst.*, **9**(1), 2002. https://doi.org/10.17705/1CAIS.00910

[18] X. Li, S. Karnan, and J. A. Chishti, "An empirical study of three PHP frameworks" in *2017 4th International Conference on Systems and Informatics (ICSAI)*, 2017. https://doi.org/10.1109/ICSAI.2017.8248546

[19] L. Lancor and S. Katha, "Analyzing PHP frameworks for use in a project-based software engineering course" in *Proceeding of the 44th ACM technical symposium on Computer science education*, 2013. https://doi.org/10.1145/2445196.2445350

[20] A. Mesbah, A. Van Deursen, and S. Lenselink, "Crawling AJAX-based web applications through dynamic analysis of user interface state changes" *ACM Trans. Web*, 2012. https://doi.org/10.1145/2109205.2109208

[21] M. Ying and J. Miller, "Refactoring legacy AJAX applications to improve the efficiency of the data exchange component" *J. Syst. Softw.*, **86**(1), 72–88, 2013. https://doi.org/10.1016/j.jss.2012.07.019

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 2, 484-493 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Interaction Model and Respect of Rules to Enhance Collaborative Brainstorming Results

Thierry Gidel[*,1], Andrea Tucker[2], Shigeru Fujita[3], Claude Moulin[4], Kenji Sugawara[3], Takuo Suganuma[5], Yuki Kaeri[6], Norio Shiratori[7]

[1]*Sorbonne Universités, Université de Technologie de Compiègne, COSTECH EA 2223, 60200, France*

[2]*Interuniversity Center for Research in Education, Université de Lille, CIREL EA 4354, France*

[3]*Faculty of Information and Computer Science, Chiba Institute of Technology, Narashino-shi, Chiba-ken, Japan*

[4]*Sorbonne Universités, Université de Technologie de Compiègne, JR Unit CNRS 7253, Heudiasyc, 60200, France*

[5]*Cyberscience Center, Tohoku University, Sendai, Miyagi, Japan*

[6]*Faculty of Media Studies, Department of Media Studies, Mejiro University,4-31-1 Nakaochiai, Shinjuku-ku, Tokyo, Japan*

[7]*Research and Development Initiative, Chuo University,1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan*

ARTICLE INFO

Article history:
Received: 14 January, 2020
Accepted: 21 March, 2020
Online: 09 April, 2020

Keywords:
Creative problem solving
Interaction model
Multi-agent system
Multi-touch device
Methodological rules

ABSTRACT

This paper presents a collaborative interaction model (CIAO) and several behaviour rules that could enhance brainstorming results. The model is composed of five elements which may be used during brainstorming sessions, consisting of activities performed by participants which are characteristic of different modes of interaction. Some sequences of these interactions may be considered more or less adequate than others. In particular, in order to get better results during brainstorming activities participants must respect certain rules when they write their ideas and when they consider notes written by somebody else. We argue that a multi-agent system can recognize different interaction modes and verify the respect of these rules by analyzing videos and notes produced by the participants in real time. Such a system must be trained as a machine learning system before being used during actual meetings. This system can simplify the role of the meeting facilitator. It can send a summary of the meeting situation, such as the proportion of each mode of interaction and identify behaviors that may need to be addressed. We present how feedback could be sent individually or addressed to the entire team. We will begin by presenting the interaction model, then propose an automatic recognition of these modes from video recordings and log analysis. We also address the necessity of rules and the structure of a multi-agent system which is able to verify whether or not the participants are respecting them. Finally, we propose how experiments could show the level of acceptance of such a system by users. The goal of this prospective research is to define a non-intrusive system that can be used during brainstorming sessions, based on an interaction model to enhance the quality of meeting results.

1 Introduction

This paper is an extension of work originally presented in the 23rd IEEE International Conference on Computer Supported Cooperative Work in Design [1]. Several research projects have proposed various architectures for supporting activities, in particular creative problem solving, in the context of collaboration between teams situated in different locations. Their front-ends typically involve rich user interfaces and some of them use large multi-touch surfaces[2]-[4]. The user interface of such tools should be synchronized between sites to ensure a continuity of experience. Their back-end is mainly based on a managing multi-agent system situated in the cloud that supports the functions of the applications (such as storing data and synchronization). To facilitate the collaboration, provide a sense of presence, and ensure awareness of other team members' actions and oral conversations, the sites must have specific equipment (cameras

[*]Corresponding Author, +33 3 44 23 52 10 & thierry.gidel@utc.fr

and microphones).

Figure 1 shows two people working during brainstorming activities around a large multi-touch table in France, at the Université de Technologie de Compiègne (UTC) while communicating with a remote team situated in Japan, at the Chiba Institute of Technology (CIT). In this image, we can see that people use virtual keyboards on the table for creating digital notes. At the same time, they can see the other team (on the right section of the board), from several angles. The team in France can also observe the similar multi-touch device used by the other team (left part of the board).

The initial stage of a project meeting is often brainstorming, which requires a facilitator to manage member contributions.

From a technological point of view, these types of collaboration environments reach a high level of quality that should allow people to obtain good results during their collaboration. However, it appears that is not always true due to some practical, methodological and behavioural reasons:

Practical reasons: people do not know how to use these systems and, for example, cannot find the right way to activate a menu, a virtual keyboard, etc. in order to perform a task. Beyond the software, the hardware may also be slow or not as responsive as the user might expect.

Methodological reasons: collaboration is a complex process with multiple steps, which are not necessarily completed in a specific order. Participants do not necessarily know which is the best step to choose, when to move to the next step or repeat previous one in order to make their collaborative activity successful.

Behavioural reasons: people do not necessarily know how to conduct themselves when they are in a group meeting. It is the role of the meeting facilitator to help, to encourage or to set limits to, for instance, ensure equal participation. Everybody should be able to participate and use competencies at their best. This can be particularly complex in multi-cultural teams with potential language barriers. Moreover, it is not always easy to find a meeting facilitator with enough expertise to manage collaborative sessions and produce interesting results.

Our long term objective is to bring an AI based support to facilitators dynamically in order to make the meeting management easier especially when several distant teams need to collaborate. We claim that such a support could be useful to tackle methodological and behavioural issues. The first step was to define an interaction model between users, presented in Section 3. The second step will be the real-time AI based analysis of all the digital footprints left by meeting participants. This will allow us to get as many indications as possible on what is happening during the meeting. Digital footprints must be understood globally: audio/video recording of sessions with several cameras and several points of view, event logs of people performing activities on the large multi-touch devices. Then, the digital footprint data needs to be analysed and we need to define how the results of this analysis are communicated to the meeting facilitator and participants.

In the following sections we present some behaviour rules concerning brainstorming session, the collaborative interaction model (CIAO), and the architecture of a framework able to deliver advice to a meeting facilitator. The last section describes how the facilitator could communicate tips and hints to meeting participants.

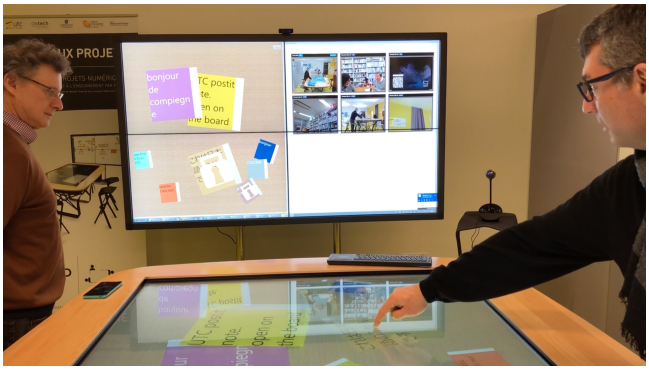


Figure 1: Environment with large multi-touch tabletop

2 Creative Problem Solving

2.1 Brainstorming

Creative problem solving is a key activity of teamwork that leads to novel ideas with great value [5]. Multi-touch and multi-user systems are expected to support this type of collaborative work by allowing cognitive stimulation through shared space [6, 7], common views, join attention, awareness [8, 9] and social interactions[10].

Group brainstorming is one of the most popular creative problem solving methods [11] but it is often misused and the role o f the group facilitator is often underestimated[12].

This seems especially critical when it comes to multi-disciplinary, multi-cultural distributed teams[13]. That is why we propose in this article to develop a complete system that could capitalize on the positive impact of multi-touch, multi-user devices based on cognitive simulation and social comparison processes[14, 15], simultaneous written and oral interactions [16, 17], AI and multi-agent based systems that would facilitate the team-work during creative problem solving and, more specifically, brainstorming sessions.

2.2 Rules

In order to encourage better results during brainstorming activities participants must follow some rules. Osborn[18] outlined four guidelines for brainstorming: (1) criticism is excluded, (2) free-wheeling and wild ideas are encouraged, (3) the greater the number of ideas, the greater the likelihood of useful ideas, (4) combination and building on others' ideas is welcome.

2.3 Barriers and propositions for effective implementation

Based on an extensive literature review, the main barriers to effective brainstorming are summarized by Isaksen [12] as (1) judgments during generation, (2) members giving up on the group and (3) an inadequate structure of interaction. In section 3 we first propose a model of interaction between users during general meetings and then we show how to analyze people's behaviour according to this model. Currently, analysis is done manually, looking at video recordings. A first objective of our research is to enable the automatic analysis of

videos, based on deep learning techniques. A second objective is to define acceptable and unacceptable structures of interaction during meetings.

To avoid judgment, which leads to inhibition due to fear of critical evaluation, voice and video analysis will track those attitudes and generate appropriate responses, either directed to one person (in a semi-private way), to a sub group (a local team on a specific site) or to the full team (simultaneously on all remote sites).

In order to avoid cognitive inertia or uniformity, wild ideas could be inserted within the group production by a computer support system in such a way as to be unable to identify the idea's origin.

One of the reasons for members giving up on the group is social loafing or freeriding. To avoid this, the setting will allow participants to view, through peripheral vision, others' production of ideas. This use of a matching of effort effect is supposed to work if there is a good group dynamic. To initiate this dynamic, the system could set challenges to stimulate the global team or the local sub-teams in terms of numbers of ideas generated, by displaying and announcing when each sub-group has produced ten more ideas. This could be associated with encouragement to extend their effort, which is known to lead to increased participation[19]. This kind of tool would require an experienced facilitator to avoid potentially harmful competition between sub-groups, or even encourage a team vs. computer competition.

System architecture and front-end software are designed to avoid inadequate interaction with devices. However, the way to favour adequate interaction between people, is still a meeting management issue. We propose a first step towards an automatic recognition of adequate and inadequate sequences of interaction between people during meetings thanks to video analysis.

The fact that all participants can write their ideas simultaneously will help to avoid production blocking [20] and limited airtime issues that are a common problem in oral brainstorming, where one person talks at a time. Simultaneous processing could even be encouraged by the challenges proposed previously. As it is true that we often think of more ideas than we could write down, a particular attention should be granted to text entry interaction: keyboard, autocompletion, word suggestion, spelling and grammar check, etc. to speed up the writing process.

We propose to develop a system that handles those major problems by helping the facilitator in his/her managing role, not aiming to replace this role. It must be seen as a critical added value. Nevertheless, the group facilitator has no ubiquitous power and he/she can only be in one location. Therefore, the system should help complete repetitive tasks and act as a relay in each sub-group. The system's contribution would include reinforcing the guidelines (follow the rules) and encouraging the participants. The system could also have a more direct interaction with the facilitator, as a timekeeper, suggesting the right time to shift from the production to the sorting phase, for instance, to maintain the group energy and productivity. That is the objective we would like to pursue.

2.4 Brainstorming process and rules

Brainstorming is a phase where people try to express ideas about the topic of a new project. Everyone is free to propose new ideas and there should be no limitations.

Generally people write ideas on post-it notes. In previous papers [2, 3] we showed that virtual post-it notes organized on large multi-touch devices present advantages over paper notes. That defines the context of our research: brainstorming activities using large multi-touch devices (tables and boards) during meetings recorded with video cameras. To respect privacy, videos are used for real time analysis as explained in the next sections. The research seeks to facilitate brainstorming by making these rules explicit.

As people are not always aware of the brainstorming rules, an automatic system supporting the respect of these social and process rules must be well integrated and as symbiotic as possible. Because it is difficult for the facilitator or group leader of a remote brainstorming session to ensure that the rules are followed in all sites, having a system supporting the animation role would be of great help. We define two types of rules: social/behavioural rules and process/note content rules.

2.5 Rules regarding behavior

During brainstorming activities, people are free to propose ideas. Nobody should feel constrained because it leads to a hesitation to express ideas. Ideas are to be analyzed in a later step, which will weed out those ideas that do not correspond to the topics developed in the project. That means that everybody must respect any proposal and not be critical of any idea during the brainstorming phase. There are at least three ways to criticize: by voice (e.g. someone saying "this is not compatible with our project"), by gesture (e.g. somebody raises an arm or turns the head or makes disapproving facial expressions) and deleting someone else's idea before discussion.

2.6 Rules regarding note content

Only the author of a note is really aware of its meaning. However, somebody else reading a note must make sense of it. So, the way to write a note must not be ambiguous. Some syntactic rules, even if they are not extensive, may avoid some misunderstandings. Notes must not contain only a few words, but complete sentences. It is a good way to ensure that people have really expressed their ideas in a way that is comprehensible for other team members. Following this rule may also help people to fully develop an idea before proposing it.

2.7 Other phases of project management

Brainstorming is usually only one step of the project or problem solving process. Ideas produced during brainstorming activities need to be organized, filtered, and sometimes repeated if the group feels that there are not enough ideas to move the project forward. For that reason, the system analyzing the respect of syntactic rules may be used for other, more semantic tasks. It should be possible for this system to recognize similar ideas, new ideas and bring attention to them. New ideas are ideas where the content has not be reused in other notes.

3 Collaborative Interaction Model

3.1 Analysis Model

Behaviors that can be observed during globally collaborative work include cooperation, individual work, presentation of that work etc. As such, we propose five modes of interaction which break globally collaborative work into its finer details: individual work, communication, coordination, cooperation and collaboration[21]. We define each of these modes as:

- Individual Work: moments when individuals retreat from the group in order to reflect and construct their ideas [22], as well as work performed on tasks with which they were entrusted by the group.
- Communication: this mode of interaction allows individuals to introduce new information into the group, creating the point of departure for a shared vision [22]. Communication can take the form of providing information orally, presentations or adding written notes into shared spaces.
- Coordination: denotes the organization of activities (events, behaviors and actions) that structure and organize tasks in order to facilitate cooperative work [23].
- Cooperation: is produced following individual work, often preceded by the division of tasks amongst group members [24]. It appears as the results are put back together. This pooling of individual work necessitates negotiation to synchronize each actors representations.
- Collaboration: designates the co-elaboration, co-evolution, or co-construction of tasks and ideas by participants in order to reach a common goal [23, 22]. The most fundamental difference between collaboration and cooperation relates to how the production is constructed: together (in the case of collaboration), to the point that it is difficult to determine who contributed what; separately (in the case of cooperation).

These modes of interactions are mobilized, in a non-linear manner, by participants during work sessions and over the course of long-term projects. The project environment, methods and tools used, intervention or instructions given by the moderator, influence the behavior of participants and as such, the emergence of these modes of interaction.

We go on to present the targeted production of the different modes as well as some elements regarding the complexity of the interactions that make them up.

3.2 Targeted Production

Targeted production leave traces either on tools or meeting video recordings. In this section we present the elements we can detect in the video.

- Individual Work: individual work is shown through reflection, aiming at the construction of ideas and meaning with the goal of eventually re-introducing the elements they judge as useful for the group. Producing written notes can be observed on video recordings and also on logs produced by the tools, because notes are written using virtual keyboards on multi-touch devices. Parsing these logs, we can also analyze the content of notes and detect those which do not respect predetermined guidelines or are interesting because the idea they contain has not yet been produced/shared with the group.
- Communication: Each member of the group is likely to have different results and ideas based on their individual work and experiences. Those ideas need to be introduced to allow new information into the group's discussions. This can be observed on recordings through speech, especially immediately following individual work or through actions on the device, such as sending a note from one surface to another (moving the idea from an individual work space to a collective one).
- Coordination: In collective sessions, coordination discussions allow for the definition of tasks and identification of responsibilities. This can be observed through the content of the conversation, related to organization and planning or through the type of tool being used, such as a GANTT. It is also typical to see a change in coordination following the intervention of a coach, as they often challenge the work that is being done.
- Cooperation: The division of tasks necessitates a pooling of work that is completed individually (or in sub-groups). This combining requires the establishment of consensus after having considered the information, opinions and arguments of each member. This can be detected based in the content of the discussion in audio recordings. Typically few notes are produced during this phase, but we see the modification of existing digital objects, instead.
- CollaboraUTCtion: The group works together to co-produce a shared vision of concepts, of solutions, strategies, which is materialized through writing, models, reports or presentations. This, again, can be detected based on the content of the discussion in audio recordings.

  The concept of globally collaborative work is introduced in [21] which examines some aspects of physical-digital workspaces, focusing on multi-user, multi-touch technologies and shows how different workspaces impact collaboration.

3.3 Manual Video Annotation

At the UTC, students have courses about project management. We installed a class environment in which there are five cubicles, each one containing a large multi-touch table and a large multi-touch board. We used software developed by the Ubikey start-up[1]. During these management courses, students work on several case studies, performing activities mediated by these new tools and experimental working methods. For research purposes, the Ubikey software allows each action on these devices to produce a digital event that

[1] Ubikey proposes a collaborative visual management tool consisting of two large multitouch surfaces (a table and a board) as well as a suite of applications which makes their use simple by supporting many kind of activities, including brainstorming.

could be tracked in near real-time on a log file. Each session produce hundreds of events registered in the log.

Each cubicle in the class environments also contain a system of four cameras able to record students' activities, behaviour and discussions. With the agreement of the students, we have recorded more than fifty hours of videos that we have manually analyzed. Figure 2 shows the view from the different cameras in one cubicle. For privacy reasons, we show the devices before the students enter the room.



Figure 2: Views from cameras for video recording

For analyzing the videos, we divided each recording into thirty seconds sequences and annotated them with one or more tags corresponding to the interaction modes described previously. This was a huge but very interesting work that consolidated our model. Figure 3 shows an excerpt of an annotation video file. For each sequence, we indicated which interaction mode appeared. The number simply represents the associated mode of interaction (8 for collaboration, 7 for cooperation, etc..) This system allowed us to avoid coding errors and graph the sequences more easily, demonstrating when and how these modes of interaction overlapped.

	A	B	C	D	E	F
	Time	Collaboration	Cooperation	Coordination	Communication	Individual Work
	00:00:00					
	00:00:30					4
	00:01:00					4
	00:01:30					
	00:02:00					
	00:02:30					
	00:03:00					4
	00:03:30				5	4
	00:04:00			6	5	4
	00:04:30				5	4
	00:05:00	8		6		4
	00:05:30				5	4
	00:06:00					4
	00:06:30					4
	00:07:00	8			5	4
	00:07:30	8				4
	00:08:00			6	5	4
	00:08:30		7		5	4
	00:09:00					4
	00:09:30		7		5	4
	00:10:00		7			4
	00:10:30		7			4

Figure 3: Excerpt of an Excel sheet containing video annotation

3.4 Automatic Video Annotation

We claim that an AI based automatic video annotation, together with a log event analysis could automate the video annotation process. Machine learning or inductive learning is based on software that learns from previous experiences and builds a model for predicting intelligent results for newly fed inputs (see figure 4). Such a computer program improves performances as more and more examples are available. In our case, we benefit of a large video corpus manually annotated. We think we can feed enough data to this machinery software for it to learn patterns, with the final output being the annotations according to the five annotation criteria defined in section 3.

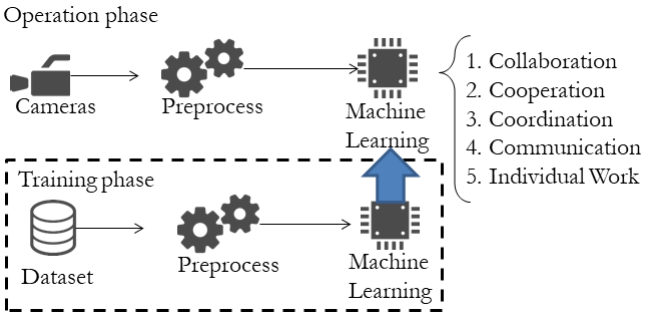


Figure 4: Video analysis, general training and prediction pipeline

Our first experiments were focused on using the already annotated video recordings to build and train a deep-learning action recognition model capable of directly recognizing each of these modes of interaction. We used the videos which were already manually annotated as training data . As of today, seven meetings have been manually annotated. Of those seven meetings, two also have the digital footprint logs associated with the devices used, and two more have been captured using four different points of view. We decided to discard the two videos taken from the ceiling viewpoint as they were too different from the rest of the dataset. This left a training dataset of around 8 GB and 12 hours. As the annotations were taken on 30 seconds intervals, we further cut the each video into the corresponding 30 seconds segments witch left us with 1473 annotated videos of 30 seconds. We finally doubled this number adding left-right flipped copies of each for a final training dataset of 2946 data points.

We based our work on the Inflated Inception model[2] described in [25] and used the TensorFlow framework [3]. This network has already been successfully used by their authors in [26] with weights trained on the Kinetics dataset [27] to improve the SOTA predictions on another action dataset performed by Google and described in [28] through transfer learning. In order to accommodate the weights, each video was first cropped to a frame size of 224 x 224 px and then had their RGB values normalized in [-1,1]. Our trials were run on a Tensorflow docker container with access to 64 Gb of RAM, Intel(R) Xeon(R) CPU E3-1230 v6 @ 3.50GHz and an Nvidia 2080 Ti with 11GB.

However, our first experiments met with some difficulties. Video sequences we annotated have a duration of thirty seconds, whereas

[2]See the following repository for the Keras implementation of this model we used: https://github.com/dlpbc/keras-kinetics-i3d

[3]TensorFlow initially developed by Google researchers is now an open source tool available in several programming languages. See https://www.tensorflow.org/

current action recognition datasets and models focus on shorter video sequences. Moreover, those datasets, and the models built upon them, focus on the recognition of atomic actions which are on a different semantic level versus the annotations we are trying to obtain. This implies the need for an intermediate semantic step in our models' architecture and the partial if not total retraining of said models. Finally, another difficulty we met was the relative limitations of our computing capabilities : the original network was trained on 500,000 iterations using 16 V100 GPUs each with at least 16GB of memory. This considerably slowed our trials with added memory and performance management, the original I3D in itself occupying 10 GB of video memory when accounting for the space needed for the input data. This is by no means a blocking factor, but hardware capabilities should not be ignored in future experimentations.

Moreover, the reason why humans choose a category when annotating is often not only based on video element recognition like a hand gesture or an interaction with an object but may also be based on vocal interactions and on their activity on the multi-touch surfaces. This led us to surmise the need of a multi modal approach combining video, audio and digital footprints. We especially believe our model will benefit from the event logs devices and thus, we plan to focus future experiments on leveraging this wealth of data to augment the video approach.

4 A Support Mechanism for Remote Brainstorming Analysis

In this section we study the questions presented above from a technological point of view and discuss the characteristics that an automatic (coaching) system must master to support brainstorming and the different tasks necessary to detect when a rule is not followed. We explain how this can be done by a multi-agent system. Each agent can detect a default and then produce an alert when something wrong appears. However some regulation is necessary in order to propose efficient feedback to participants. Also, when rules are correctly followed and the team is producing good results, we detail how this could be encouraged.

4.1 Architecture

We are able to make use of existing system architectures, which have been designed to support team activities (managing support system). Figure 5 shows a simplified architecture of a management system (left) and the added coaching system we plan to develop (right). Notes and video streams are sent by the local devices to the cloud managing system, where they are treated by its agents and dispatched to the other sites.

We will design another multi-agent system (coaching support system) which can be installed on the same cloud server that receives the streams from the different sites thanks to the managing support system agents; it will produce feedback that can be dispatched to the sites by the managing system agents. The new system can receive the streams because it can request them from the managing agents. It will be necessary to develop a new interface to provide feedback to each site, but, from a technical point of view, the work required to design and implement it is negligible. A multi-agent system is suitable for developing a heterogeneous distributed system. Our multi-agent system is in compliance with the FIPA architecture and uses the FIPA-ACL communication protocol. Agents can make decisions according to a rule-based inference system. Rule descriptions are declarative and rules are associated with a dynamic priority level. Agents run procedures associated with the rules that are chosen. No machine learning method is implemented here.

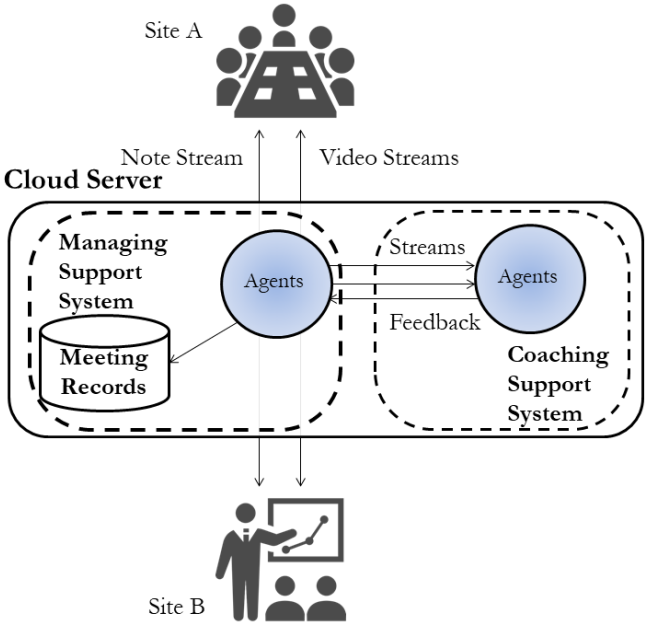


Figure 5: Multi-agent systems managing collaboration between remote teams

We believe that the coaching system must consist of two layers: a detection layer and a regulation layer. In the first layer, the agents detect the activities done by each person (production of an idea, criticism, facial expression...) and produces alerts after each detection. Other agents track the number of ideas produced during a session and produces alerts based on a comparison between subgroup metrics (i.e. returning to the notion of competition), standard production metrics (i.e. the systems default settings for what good production looks like), and on group goals and/or the specificity of the group according to the facilitator.

In the second layer, regulation agents analyze the different alerts (positive and negative) coming from the detection layer agents. According to pre-defined rules, these agents decide what feedback should be given, when and to whom (see figure 6).

4.2 Identification of participants

It is necessary to identify all the participants in a meeting. An identifier is sufficient, personal data should not be required. The system must give feedback either to one participant, to a sub-group at one specific site or to all the participants at all sites. It is when a message has to be sent to one individual or to identify the source of a particular idea that the identifier is necessary.

When using a large multi-touch surface, people use virtual keyboards to enter content into post-it notes. It is easy to add a unique

symbol on this keyboard and then a camera can associate a participant to a keyboard for identification. If necessary, facial recognition maybe used to add identifiers to participants. An piece of individual feedback, dedicated to a single participant, could be achieved thanks to an individual written notification, like a pop-up attached to the virtual keyboard he/she is using, making it discreet.

4.3 Behavioural Rules

According to the examples we proposed in the previous section, there are three ways to express disapprobation of an idea: by voice, by gesture or by deleting a note. Because of this, we believe that the installation must include different types of devices: i) cameras for video streams transfer, ii) microphones for better capturing individual voices and iii) multi-touch screens where the activity is performed. Several video streams are transferred to agents belonging to the detection layer. Agents use their own algorithm to analyze either gestures, voice or faces. Representations of notes are also transferred to detection layer agents.

4.4 Syntactic Rules

A note contains a text which expresses an idea about the brainstorming subject. The rule concerning an idea specifies that the text must be a complete and expressive phrase. This avoids any misunderstanding and ambiguity. The analysis of a note is described in section 4.6.

In the detection layer an agent is associated with each participant (i.e. with each identifier). It receives structures representing notes containing their content, the participants' identifiers and other metadata (timestamp, etc.). The natural language processing support of each agent is able to detect acceptable notes and refuse others.

The system uses different kinds of rules. Rules will be checked through experimentation and inconsistencies, then adjusted manually.

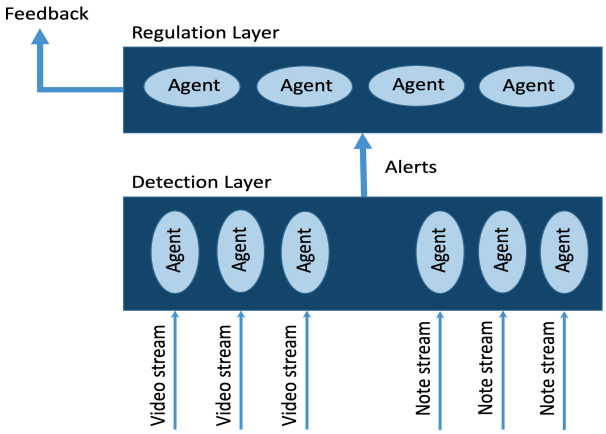


Figure 6: Multi-agents system layers

4.5 Regulation System

The regulation agents consist of two main categories: behaviour agents that manage behavioural alerts and regulation agents that manage syntactic alerts. All the alerts produced during the activities are compiled in shared memory (see figure 7). Globally, the regulation agents have different tasks: they determine the given situation of the meeting: is it the beginning, middle or end of the session?; they count the number of positive alerts (rules followed, production rate) and negative alerts (rules not followed); they check the production rate of each participant (a participant is considered inactive when the number of notes is low); they propose feedback.

A feedback agent is in charge of ordering and producing definitive feedback to participants. Feedback agreement from participants (a record of whether or not the feedback is acceptable or helpful) are also stored in the shared memory (however their exploitation will be performed in a future research).

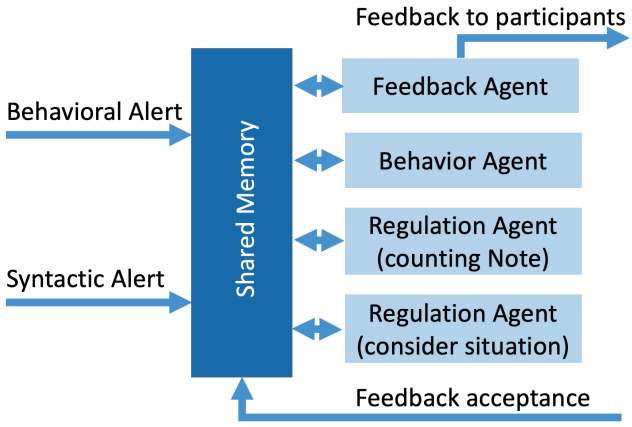


Figure 7: Regulation layer

4.6 Design of Text Processing System

This subsection describes the life cycle of a note within the agent layers (see figure 8). The note counting mechanism counts each participant's note. The number of notes may launch feedback that is designed to reinforce his/her enthusiasm.

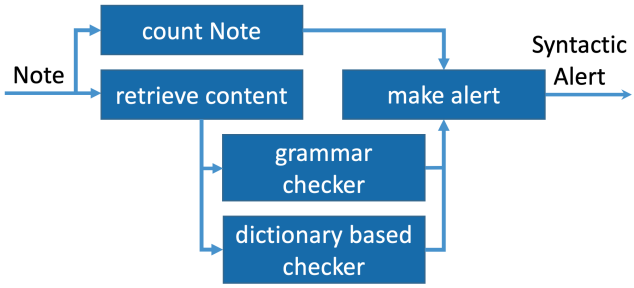


Figure 8: Text checker on note

The text of a note is checked by a dictionary and grammar checker. The dictionary based checker verifies if words are contained in predefined dictionaries and associates a value to each word. The grammar checker verifies syntactically the text of each note using a tool such as the LanguageTools. An alert is launched if this global evaluation is not high enough.

4.7 Design of Image and Voice Processing System

In this subsection we detail the main components of a checking mechanism for a video stream (see Figure9): The facial recognition mechanism identifies a participant; the gesture recognition mechanism categorizes a participant's actions; the emotion recognition mechanism associates a participant's face to a class of emotions; finally the tone of voice recognition mechanism associates a participant with a class of emotions. The voice recognition mechanism achieves a speech to text transformation.

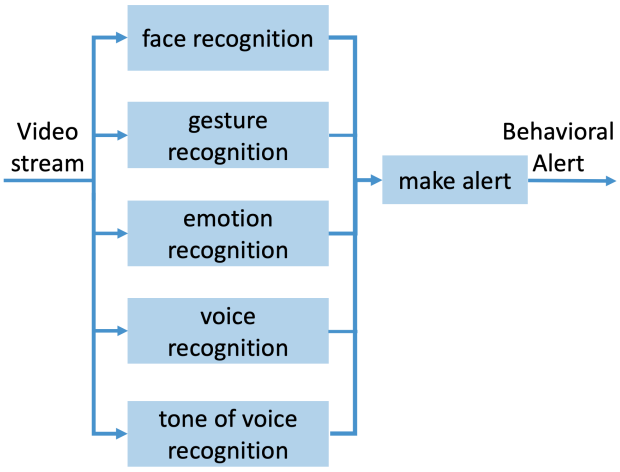


Figure 9: Video checker

These recognition systems require large data sets to learn to categorize faces, emotions and gestures. Alerts produced by this mechanism will have a relative score which indicates the level of reliability of the identification.

Different tools can be used for these mechanisms. Open source projects OpenCV[29] with YOLO[30] are suitable for identifying a participant. The OpenPose tool[31, 32] can detect human body, hands, facial and foot key-points on single images; the Microsoft Emotion engine[33] recognizes emotion categories from Kinect sensors.

Of course, non-verbal communication depends largely on participants' language, nation, region, age or personal experience and its analysis requires a good deal of testing.

5 Feedback For Rule Respect

In this section we detail the way an automatic system could give feedback in real time to the participants of a brainstorming meeting about the respect of rules. The feedback is based on the different types of rules we identified in previous sections.

The main questions are when, how,to whom to deliver the feedback and if hints are necessary. Another question is how to avoid interfering with the meeting by being as discreet as possible, while also being sensitive to participants' emotions to help create a safe psychological space.

Both individual and generic messages have to be integrated in the coaching system. Individual notifications are dedicated to a specific person, for instance because he/she does not follow one rule consistently. This means that the system is able to refer to and recognize all the participants of a meeting. This could be achieved through an individual written notification, like a pop-up attached to a keyboard (see section 4.2).

Common notifications can be sent to a group of people, for instance, a local team (i.e. a distant team would not receive it). If the regulation layer of the coaching system considers that the same rule is not followed by several people in the same place, a local verbal or written message could be sent to the local team. A verbal message could be heard by the other team, so if we want to avoid this possibility, it is better to choose to send a message written on the large multi-touch surface.

Social interactions are important in brainstorming activities. The notifications should not divert participants from their work. Therefore, the number of notifications should be adjusted to the situation and should take into account the group dynamic and notification content.

The balance between feedback to motivate or correct behaviors should be taken into account. This means positive comments such as "good job, you are producing lots of ideas, keep it up!" are as important as constructive feedback such as "be careful, you should not criticize other people's ideas".

Finally, specificity of the team should be taken into account in the way the feedback is delivered. For instance, cultural bias could lead to different ways of producing and interpreting feedback.

6 Evaluation Scenario

In this section we explain how we will test the coaching system. It is necessary to evaluate the efficiency of the system at both a technical and acceptance level.

The first step is to test the quality of the coaching system at the detection layer with the production of alerts. Different scenarios creating notes automatically will be played and statistics about the number of alerts delivered by the system analyzed. Video with specific gestures and faces will be created and proposed to the agents. When the rate of recognition is satisfactory we can begin the second step of our evaluation, i.e. the test of the regulation layer.

Based on pre-recorded videos and notes from a session, we determine what feedback individuals and the team should receive. It is a test relying on subjective judgments made by researchers, but based on the rules outlined previously. Researchers will manually balance individual and global feedback delivery for timeliness and content. When a satisfactory rate of feedback is delivered by the coaching system we can begin the third step of the evaluation, i.e. experiments with real scenarios and real participants playing foreseen sequences.

The last step of the experiment will focus on the way people accept to respect constraints during brainstorming activities and receive feedback during their work. Prior to the experiment, participants will receive instructions about the rules to be followed during the session in order to make the conditions explicit. We will conduct experiments in two ways: i) brainstorming activities managed only by a human in charge of the facilitation to ensure the respect of the rules, ii) brainstorming activities where an automatic system gives feedback together with a human facilitator.

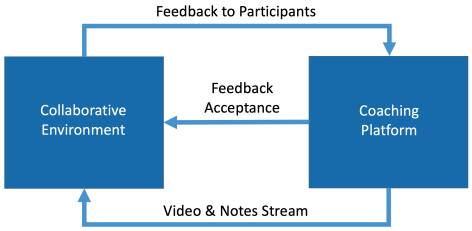


Figure 10: Exchanges between systems

The results of the last experiment will be evaluated by observers who monitor how people react when they receive feedback. We postulate that the user interface may allow participants to react when they are notified, with a very quick action such as pressing a button on the notification itself. In this way, we can consider a loop where the coaching system registers the level of feedback acceptance by the participants (see figure 10).

Our objective is not to cancel out the role of the facilitator because integrating a coaching system into pre-existing support system architecture can be difficult, but also because the role of the facilitator is fundamental. The coaching system is only an assistant which allows the facilitators to devote time to their main tasks, such as managing the group dynamic.

7 Research limitations

This position paper defines our research goals, details our propositions on how to achieve this goal and describes our three step research protocol for evaluation. As presented, we are already facing difficult technical challenges that we need to overcome in order to carry out this research, such as automatic video annotation issues. We also foresee some tricky problem to be solved with regards to fine tuning the rules implemented in the regulation system. Finally, once the technical issues are be solved, acceptance of this type of system by both the participants and the coach needs to be thoroughly evaluated.

8 Conclusion

In this paper, we have presented how brainstorming results could be enhanced if meeting participants observed several simple practical, methodological and behavioural rules. We have proposed a model of interactions between people during collaborative meetings. We have described how an AI based analysis system could both summarize the state of a meeting according to this model and give advice to a meeting facilitator and the participants.

We hypothesise that a technological support system could help meeting facilitators with their group management tasks if it is able to analyze whether or not behavioural and methodological rules are followed and if it produces efficient and adapted feedback.

We proposed an architecture for such a system and explained the way we can conduct the necessary experiments for testing its implementation and the way people receive and agree with the feedback.

In the context several teams situated in different locations, this model and associated architecture could help supporting activities and in particular creative problem solving. It therefore could be integrated into existing computer supported collaboration tools along with video-conferencing systems. In a future work, we would like to implement new technical functionalities such as an automatic learning system based on analysis of previously recorded sessions. Additionally, we would like to study the retroactive feedback acceptance by participants on the system itself. For instance, an agent would be able to understand and adjust it's behavior in order to achieve the highest impact on the targeted participant. This mean that the agent would need to learn from participant feedback acceptance and choose the appropriate way to give feedback to each participant.

References

[1] T. Gidel, S. Fujita, C. Moulin, K. Sugawara, T. Suganuma, Y. Kaeri, and N. Shiratori. Enforcing methodological rules during collaborative brainstorming to enhance results. In *2019 IEEE 23rd International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pages 356–361, 2019.

[2] Y. Kaeri, K. Sugawara, C. Moulin, and T. Gidel. Agent-based design of iot applications for remote brainstorming support. *2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design ((CSCWD))*, pages 820–825, 2018.

[3] Y. Kaeri, C. Moulin, K. Sugawara, and Y. Manabe. Agent-based system architecture supporting remote collaboration via an internet of multimedia things approach. *IEEE Access*, 6:17067–17079, 2018.

[4] Y. Kaeri, K. Sugawara, C. Moulin, and T. Gidel. Agent-based management of support systems for distributed brainstorming. *Advanced Engineering Informatics*, 44:101050, 2020.

[5] *The International Handbook of Creativity*. Cambridge University Press, 2006.

[6] P. Marshall, R. Morris, Y. Rogers, S. Kreitmayer, and M. Davies. Rethinking 'multi-user': An in-the-wild study of how groups approach a walk-up-and-use tabletop interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 3033–3042, New York, NY, USA, 2011. ACM.

[7] Y. Rogers, Y. kyung Lim, W. R. H. PhD, and P. Marshall. Equal opportunities: Do shareable interfaces promote more group participation than single user displays? *HumanComputer Interaction*, 24(1-2):79–116, 2009.

[8] R. Fleck, Y. Rogers, N. Yuill, P. Marshall, A. Carr, J. Rick, and V. Bonnett. Actions speak loudly with words: unpacking collaboration around the table. In *Proceedings of the ACM international conference on interactive tabletops and surfaces*, pages 189–196. ACM, 2009.

[9] S. Conversy, H. Gaspard-Boulinc, S. Chatty, S. Valès, C. Dupré, and C. Ollagnon. Supporting air traffic control collaboration with a tabletop system. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*, CSCW '11, pages 425–434, New York, NY, USA, 2011. ACM.

[10] O. Shaer, M. Strait, C. Valdes, T. Feng, M. Lintz, and H. Wang. Enhancing genomic learning through tabletop interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2817–2826, New York, NY, USA, 2011. ACM.

[11] K. L. Dugosh and P. B. Paulus. Cognitive and social comparison processes in brainstorming. *Journal of Experimental Social Psychology*, 41(3):313 – 320, 2005.

[12] S. G. Isaksen and J. P. Gaulin. A reexamination of brainstorming research: Implications for research and practice. *Gifted Child Quarterly*, 49(4):315–329, 2005.

[13] N. Michinov and C. Primois. Improving productivity and creativity in online groups through social comparison process: New evidence for asynchronous electronic brainstorming. *Computers in Human Behavior*, 21(1):11 – 28, 2005.

[14] B. A. Nijstad, W. Stroebe, and H. F. Lodewijkx. Cognitive stimulation and interference in groups: Exposure effects in an idea generation task. *Journal of Experimental Social Psychology*, 38(6):535 – 544, 2002.

[15] K. Dugosh, P. Paulus, E. Roland, and H.-C. Yang. Cognitive stimulation in brainstorming. *Journal of Personality and Social Psychology*, 79(5):722–735, 2000.

[16] B. A. Nijstad, W. Stroebe, and H. F. Lodewijkx. Production blocking and idea generation: Does blocking interfere with cognitive processes? *Journal of Experimental Social Psychology*, 39(6):531 – 548, 2003.

[17] M. Diehl and W. Stroebe. Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology*, pages 497–509, 1987.

[18] A. F. Osborn. *Applied imagination; principles and procedures of creative thinking*. Scribner, 1953.

[19] D. Johnson and R. Johnson. *Learning together and alone: cooperation, competition, and individualization*. Prentice-Hall, 1975.

[20] C. M. Hymes and G. M. Olson. Unblocking brainstorming through the use of a simple group editor. In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work*, CSCW 92, page 99106, New York, NY, USA, 1992. Association for Computing Machinery.

[21] A. Tucker, T. Gidel, and C. Fluckiger. Designing physical-digital workspaces to support globally collaborative work. In *Proceedings of the Design Society: International Conference on Engineering Design*, pages 109–118. Cambridge University Press, 2019.

[22] S. D. Teasley and J. Roschelle. Constructing a joint problem space: The computer as a tool for sharing knowledge. In S. P. L. . S. J. Derry, editor, *Computers as Cognitive Tools*, pages 229–258. Hillsdale, NJ: Lawrence Erlbaum Associates, 1993.

[23] M. J. Baker. Collaboration in collaborative learning. *Interaction Studies*, 16(3):451–473, 2015.

[24] K. A. Bruffee. Sharing our toys: Cooperative learning versus collaborative learning. *Change: The Magazine of Higher Learning*, 27(1):12–18, 1995.

[25] J. Carreira and A. Zisserman. Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset. *arXiv:1705.07750 [cs]*, Feb. 2018.

[26] R. Girdhar, J. Carreira, C. Doersch, and A. Zisserman. A Better Baseline for AVA. *arXiv:1807.10066 [cs]*, July 2018.

[27] W. Kay, J. Carreira, K. Simonyan, B. Zhang, C. Hillier, S. Vijayanarasimhan, F. Viola, T. Green, T. Back, P. Natsev, M. Suleyman, and A. Zisserman. The Kinetics Human Action Video Dataset. *arXiv:1705.06950 [cs]*, May 2017.

[28] C. Gu, C. Sun, D. A. Ross, C. Vondrick, C. Pantofaru, Y. Li, S. Vijayanarasimhan, G. Toderici, S. Ricco, R. Sukthankar, C. Schmid, and J. Malik. AVA: A Video Dataset of Spatio-temporally Localized Atomic Visual Actions. *arXiv:1705.08421 [cs]*, May 2017.

[29] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

[30] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 779–788, June 2016.

[31] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *CVPR*, 2017.

[32] T. Simon, H. Joo, I. Matthews, and Y. Sheikh. Hand keypoint detection in single images using multiview bootstrapping. In *CVPR*, 2017.

[33] Ms emotion engine. https://azure.microsoft.com/en-gb/services/cognitive-services/emotion/.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 3, 493-500 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Smart Transmission Line Maintenance and Inspection using Mobile Robots

Thongchai Disyadej[1,*], Surat Kwanmuang[2], Paisarn Muneesawang[3], Jatuporn Promjan[1], Kanyuta Poochinapan[4,5]

[1]*Transmission Line Maintenance Department, Electricity Generating Authority of Thailand, 50300, Thailand*

[2]*Department of Mechanical Engineering, Chulalongkorn University, 10330, Thailand*

[3]*Department of Electrical and Computer Engineering, Naresuan University, 65000, Thailand*

[4]*Centre of Excellence in Mathematics, CHE, Si Ayutthaya Rd., Bangkok, 10400, Thailand*

[5]*Advanced Research Center for Computational Simulation, Chiang Mai University, 50200, Thailand*

ARTICLE INFO

Article history:
Received: 21 January, 2020
Accepted: 07 June, 2020
Online: 21 June, 2020

Keywords:
Transmission line
Maintenance
Robot

ABSTRACT

The paper presents sharing several of experiences and practices on smart robotic application for overhead transmission line maintenance and inspection. First, the pilot-line pulling robot is an invention used to pull a pilot line which is an important step for additional high voltage conductor installation. The puller robot can traverse the overhead ground wire, OHGW, and pulls a lead line via a set of cradle blocks at intervals, carrying the line above the ground. The robotic puller passes over barriers below the power line, such as the road with traffic, power distribution lines, river, or vegetation making tasks achieved conveniently, safely, and rapidly without impact on nearby communities. The robot was further utilized to pull a lead line/conductor crossing over the electrical substation without interrupting energy and pull a lead line for the improvement of transmission line ground clearances. The developed pilot-line pulling robot has been accredited as the corporate best practice; the standards for innovation, operation, and maintenance are archived for works at all EGAT transmission line operation & maintenance units nationally. Moreover, EGAT was now jointly investigating with universities on a new robotic device for aerial transmission line inspection. The target of the research is to create a mobile robot prototype for inspection of overhead power lines. The inspection robot shall crawl along the ground wire and transpose autonomously across installed equipment on the ground wire, such as vibration dampers, suspension clamps, compression dead ends, etc. In addition, the inspection robot is able to take photos and videos during a transmission line inspection in both offline and online features. Using the robot, transmission line inspection's labor cost can be reduced, and the new method helps improve patrol and inspection efficiency, comparing to the conventional manpower method. Trough utilization of the new maintenance and inspection robots, utilities can minimize transmission line operation & maintenance budget.

1. Introduction

Electricity Generating Authority of Thailand (EGAT) is Thailand's leading state enterprise about electric energy under the supervision of the department of energy that has its main business in generation, transmission, and energy sales to other power distribution utilities. In order to promote the development of the national economic and industrial field in conjunction with the "Power for Thai Happiness" corporate agenda, the electricity utility's mission is to preserve the efficiency of the transmission network for the entire country. In the view of electrical providers, the primary purpose of high voltage transmission systems is the provision of services. EGAT also has a long tradition of

*Thongchai Disyadej, EGAT, Tel. +66 53 220048, Thongchai.d@egat.co.th

developments in the service and management of the energy supply sector utilizing the robotic technology, as demonstrated in Figure 1. Northern Region Operation Division developted an Optical Ground Wire (OPGW) inspection with the robot for examining a damage spot on OPGW's strands in 2008 [1]. The OPGW inspection robot was upgraded to a better version, 2 years later, such that phase conductors can be inspected also. In 2014, Electrical Maintenance Division and a domestic university co-invented a robot prototype used to inspect boiler water-wall tubes [2]. Utilizing magnetic adhesion, the NDT inspection robot can examine the thickness and the external properties of boiler water-wall tubes. Furthermore, Electrical Maintenance Division funded the national technology center on the same year a project on a tiny robot prototype with 2-cm size or less for internal generator examination [3]. Similar to other robots on the same sector, owing to their smaller form, the robot may analyze inside the generator a distance of less than 3.5 cm. This saves generator testing period and expensive oversea testing robots. In addition, EGAT funded the national metal and materials technology center an R&D work in 2013 about a robotic system development for welding [4]. Main findings include a welding device with applications for welding process monitoring and torch motion creation for overlay welding of parts used in power plants. Afterwards, in 2017, another development for robotic welding system to be used in the repair phase for high-value power plant components was performed [5]. Recently developed [6], new transmission line maintenance and inspection robots are proposed in this paper.

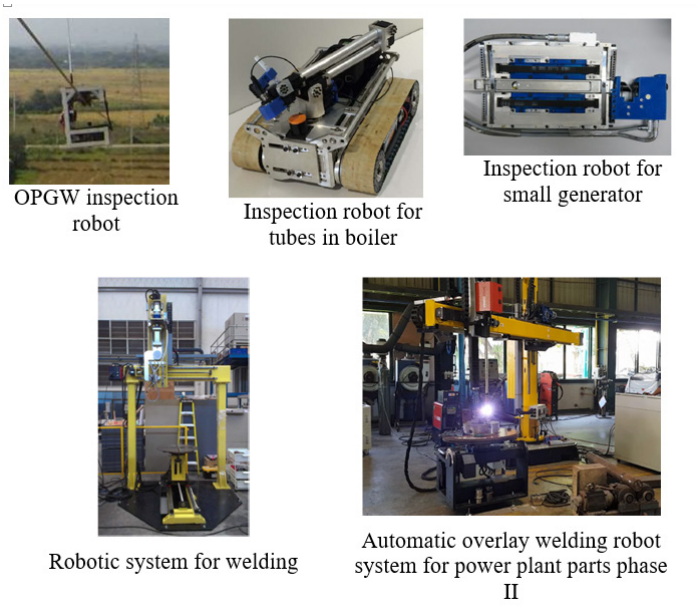


Figure 1: Examples of EGAT innovation on robots [1-5]

A summarized literature review is performed on transmission line maintenance and inspection robotics developed by many researchers. LineScout is a mobile robot developed by the Hydro-Quebec Research Institute to perform live line inspection and maintenance [7-15]. Expliner is an inspection robot developed by the HiBot Corporation, Kansai Electric Power Corporation, and Tokyo Institute of Technology for commercialization [16-17]. Also, TI is an inspection robot developed, evaluated, and demonstrated by Electric Power Research Institute under EHV condition [18]. Moreover, various robot prototypes focusing on 220-500 kV transmission lines are developed by Chinese academy of sciences and many Chinese universities [19-22]. Three types of high voltage transmission line robots are developed by Wuhan University for single and bundled conductors [23-30].

According to presented relevant works, many robots were designed with some limitations for actual field working, e.g. their weight and dimension, specific configuration of conductors, limitation for some tower junction, difficulty for actual field works, and etc. Therefore, two transmission line maintenance and inspection robots with light weight, small size, and ease of use are designed and developed in order to be a new realistic solution for power utilities expecting more efficient O&M system.

The paper sections are arranged as the following. We present in section 2 a developed transmission line maintenance robot. We firstly summarizes problems and constraints of the existing pilot-line pulling method for additional stringing of conductor projects. Then, the development of the new maintenance robot and actual success implementations in transmission lines are presented. The next part, section 3, deals with a novel inspection robot for transmission lines. The background, design, Lab test, and field test of the inspection robot are presented. Working results of both developed robots are discussed in the next section. Evaluation and future improvement are finally concluded in section 5.

2. Transmission Line Maintenance Robot

According to the corporate social responsibility policy, transmission system division desires to improve a better approach for stringing additional power lines while communities suffer no public impacts. Thus, the team of inventors decides to evolve a research project on an advanced approach of pilot-line pulling for the power line installation that conventionally requires manpower and land/water vehicles to pull out a pilot line dragged over the field beneath the power line above, shown as Figure 2-3. The consequence is that it damages farmlands beneath the power line and causes an effect on communities, Figure 3.

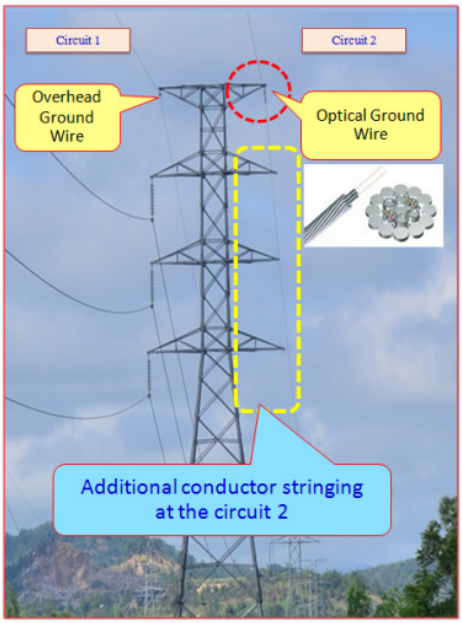


Figure 2: Conductor stringing for the circuit no. 2 to be executed on the blank crossarms of towers

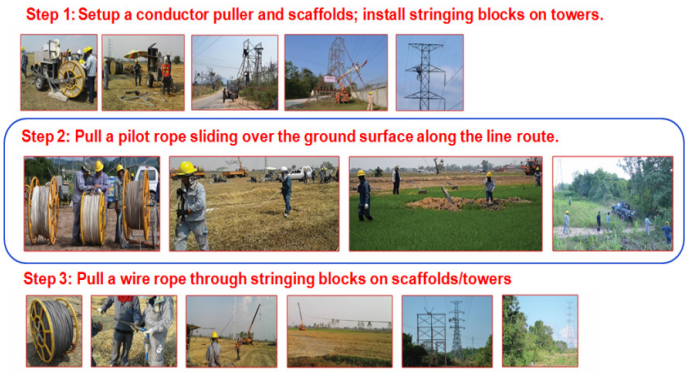


Figure 3: Conventional steps for pilot-line pulling



Figure 4: Constraints and challenges on pilot-line pulling

Constraints and challenges for pulling the pilot line with the existing working approach are concluded as the following:

- Interrupted traffic and vegetation damages to landowner (damage compensation to be paid)
- A great amount of workers needed co-working on different places
- Risks of injury by incidents during line pulling operations over the roads, railways, communication lines, power distribution lines, and etc.
- Scaffolds needed above the street and electric distribution line
- Electricity de-energization required when crossing over high voltage switchyards

Therefore, the researchers utilize the robotic technology to diagnose and evolve the new creative approach for better pilot-line pulling operation above all obstructions with the robot developed as a tool to pull out a pilot line through a number of cradle blocks. The complete description of the specially designed puller robot together with a series of cradle blocks is shown in Figure 5. Whilst pulling, the running line is carried on the OHGW at structures and lifts at the height more than 13 m. over the ground, in Figure 6. Thereby, communities suffer no more problems from our jobs, and damage compensation for properties does not need to be paid anymore. Moreover, construction workers can work on stringing a lead line above the highway, river, and live power line with safety, efficiency, and reliability. Additionally, the robot has been utilized further for a lead-line pulling over the live substation in the absence of service outage and also a lead-line pulling for the ground-clearance upgrading of an overhead power line.

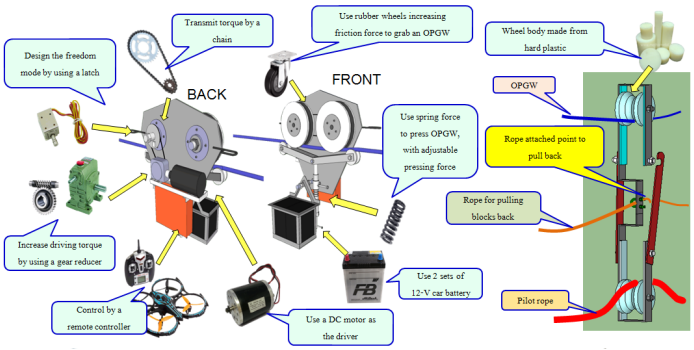


Figure 5: The composition of robot for line pulling work

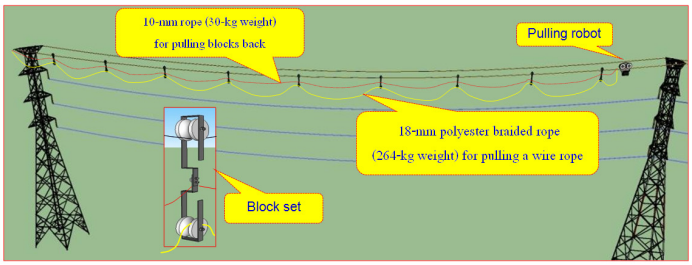


Figure 6: The complete description of the new pilot-line pulling method



Figure 7: Actual implementations of the transmission line maintenance robot

With success implementation, EGAT has innovated a special robot for pilot-line pulling over all obstructions, and it passed the pulling-force test with 60-kilogram load and 35-degree inclination. The consequences have proved the accomplishment.

3. Transmission Line Inspection Robot

3.1. Background

Throughout the world, inspection of transmission lines usually consumes workforces to achieve the routine tasks by ground patrolling along the transmission line paths. Minimization of man-hour is necessary in order to improve capability while the precision of data is still needed to be conserved. For this reason, partially/completely utilizing robotic technology into the standard line- inspection working method is of great essential. Existing means of routine power line inspection in the past have been dealing with both helicopters and UAVs to minimize cost of manpower. However, the former demands costly capital and operation budget whilst the latter suffers with many constraints, such as short battery lifetime, relevant laws, dependability in operation, and so forth. Therefore, application of an innovative inspection robot rolling along an OHGW can play a significant role in smart transmission line inspection improvement for electric utilities.

The maintenance robot in the earlier section that travels along an OHGW can be practically upgraded to an inspection robot. Nevertheless, the service robot is unable to cross through sequences of obstacles, e.g. vibration dampers, suspension clamps, compression dead ends, and other hardware. That is, the robot has ability to travel just one single span between towers only. To significantly improve inspection efficiency, a new autonomous robot negotiating and crossing successive obstacles automatically must be developed.

Crossing through each tower top is the key challenge for the robotic transmission line inspection. When the robot encounters a tower, personnel needs to climb to each tower and then manually transpose the robot to another side of span. It will be very much more efficient to use if the inspection robot can automatically navigate through tower tops. Nevertheless, autonomous crossing of obstacles repetitively requires huge efforts on advanced technology because equipment installed on the transmission structures have a lot of variety. Regarding mentioned constraints, the research team has designed an approach by developing a new mobile robot which can travel through a series of obstacles and earth wire crossarms with automatic navigation as well as feature of gathering high-quality visual information, e.g., damaged equipment and right-of-way vegetation data [31].

The aim of this research project is to create a novel transmission line section robot that can roll along an OHGW with autonomous crossing of obstacles. This working approach improves more efficiency than the existing approaches, such as ground patrols or helicopters. Photo and VDO via both offline and online (real time) modes can be recorded by the robot; it is a new mean to improve more accurate inspection than the conventional mean. In addition, for an outage event, the robot can be sent out quickly to search for an outage cause in mountainous regions, minimizing the customer's service interruption.

3.2. Conceptual design of transmission line inspection robot

We designed an overhead power line inspection robot to work with autonomous feature and use a communication network for teleoperation. Figure 8 presents the conceptual design of the inspection robot. One can observe that the control portion is divided into 3 key portions systematically connected inside: moving control, communication control, and remote control parts. Those three units link altogether operations for the entire system and link with operators through an App on a handheld equipment in the field that will track data of the robot transmitted via the cloud network system, which is illustrated in Figure 8.

- Robot motion-control part

Robot movement is controlled by a microprocessor inside, which is the core duty for moving towards a desired location precisely. The inner subsystem includes a system of battery power supply which will energize power to all parts and a feedback-control part that can work with a communication-control part closely for sensing the state of the robot such that the microprocessor can make a judgement to receiving information to command motion of each movement.

- Communication control part

This part includes an embedded microprocessor which has a role about communicating to each equipment, e.g. sensors which transmit feedback information to the movement part that identify the movement state or a camera that record images during the robot roll along each path.

- Tele-monitoring control part

The robot has a communication unit with a network for transmitting states back to the cloud system via a mobile cellular system for a tele-monitoring control part. M2M (machine to machine), which is a protocol commonly for the IoT world, is used as a format of Data streaming. It can work with reliability although the restriction of system bandwidth, i.e. sending data in the field to the cloud system via the App for observing.

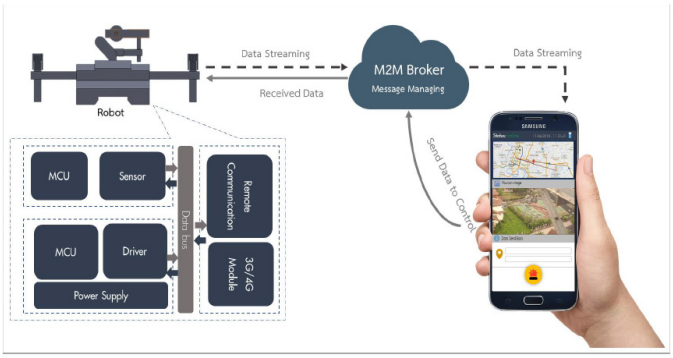


Figure 8: The conceptual design of the transmission line inspection robot

3.3. Purposed design

Figure 9 illustrates a design concept of an efficient robotic prototype for transmission line inspection with governing mechanism to transpose through obstacles on transmission lines. The model is designed and simulated using a special computer software. The robot consists of 6 linear actuators with the

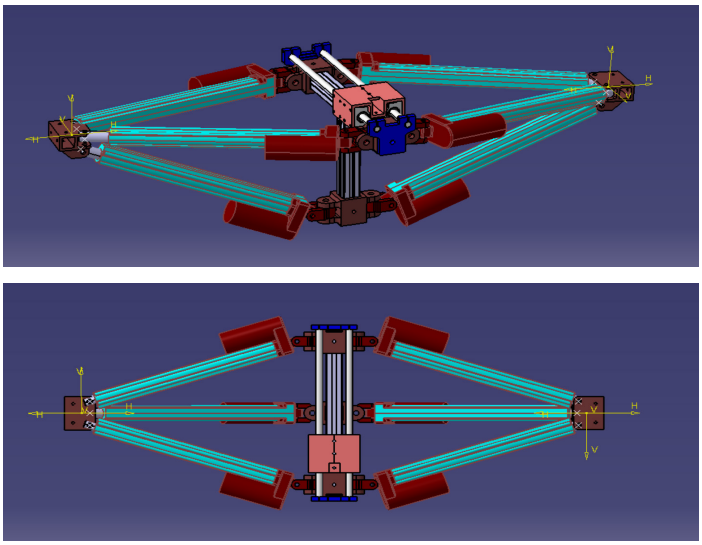


Figure 9: Design concept of the inspection robot

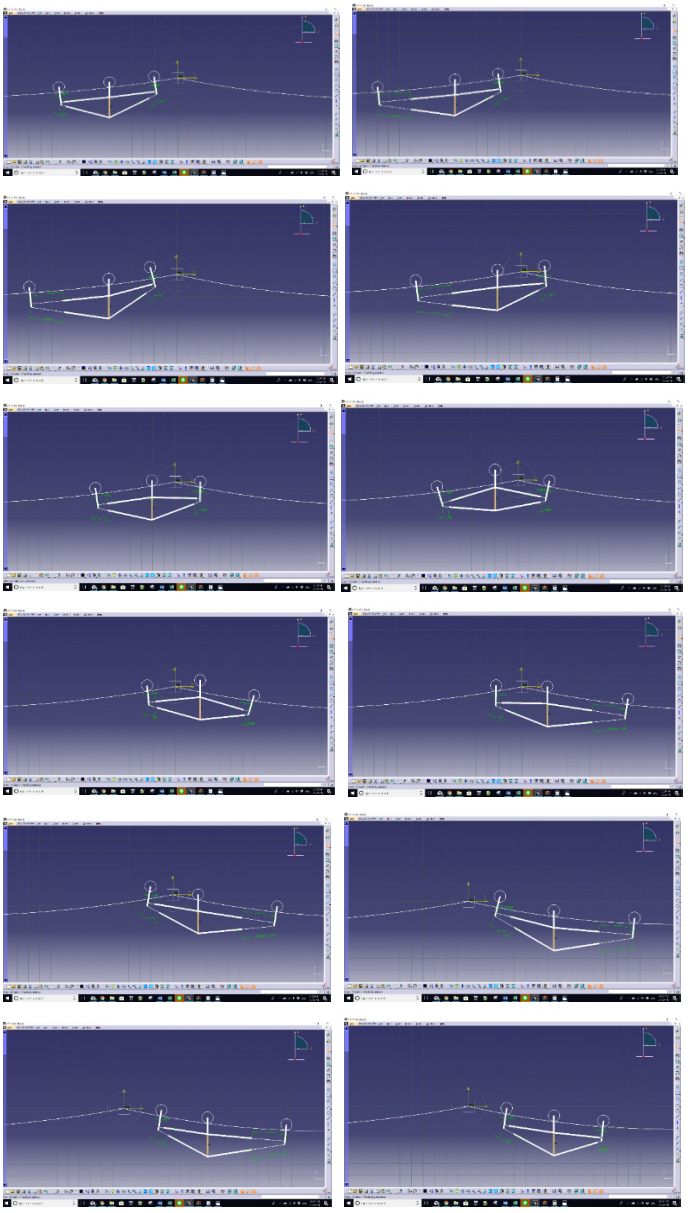


Figure 10: Transposition steps of the robot

maximum extension of 20 cm. They are used to control the front or back wheel towards the desired position. By connect 3 linear actuators together to from a parallel mechanism, the resulted actuator has very high stiffness and compact while the end-effector is still able to reach in required envelop. Moreover, there are 3 wheels for the driving of the robot. Each of wheel is installed with a driving motor which is a 200W BLDC hub motor installed inside the driving wheel set.

3.4. Motion of the robot and obstacles mitigation

The robot can shift its center of gravity to the front and rear by extending its arm to the desired direction. This enables the robot to lift the other arm out of the OHGW without losing stability.

The motion process is that the robot usually run along the OHGW and can transpose through obstacles, e.g. equipment and hardware on the structure with steps as the following:

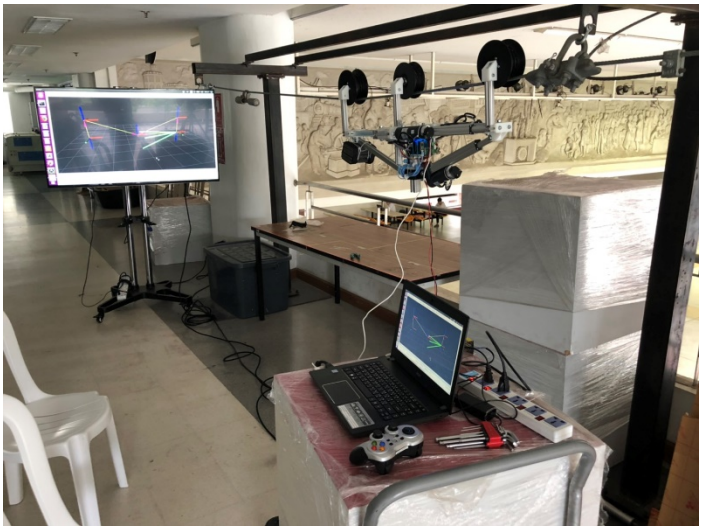


Figure 11: Laboratory test for the prototype with the line model

3.5. Kinematic model

The robot front and back arms are symmetric and identical. Each arm has three linear actuators arranged in the form of a tripod. The motion of the end effector can be described using coordinate system in the Figure 12.

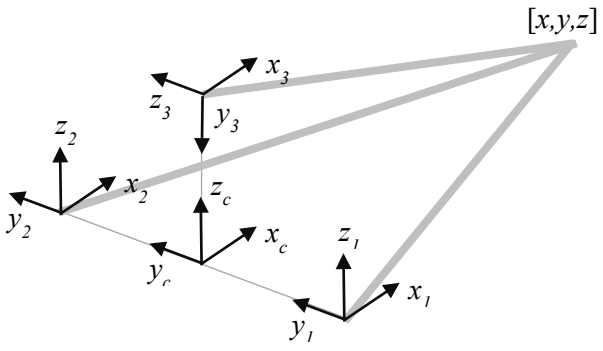
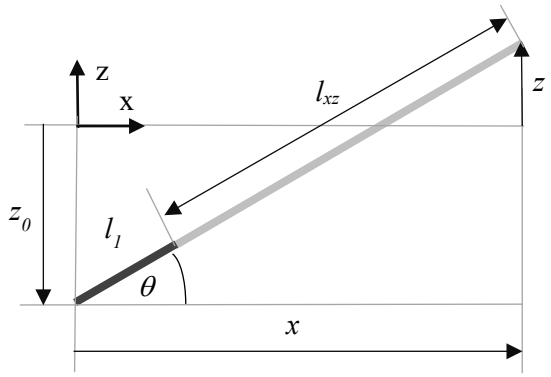
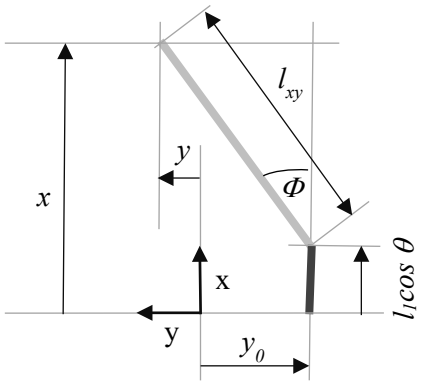


Figure 12: Coordinate system for each actuator and arm's center and also the end-effector position.

The inverse kinematic model of each actuator is an actuator length (l) as a function of the end-effector position in arm's center frame (x, y, z). There is a restriction in designing the end joint of the actuator which results in offsets between y and z rotation axis (l_1). Since all actuators are identical, we can derive the model for all actuator in each actuator's coordinate frame in the following:



a) Side view



b) Top view

Figure 13: a) Side view and b) Top view of the mechanical model of a single linear actuator assembly

$$\theta = \tan^{-1}\left(\frac{z-z_0}{x}\right) \quad (1)$$

$$l_{xy}{}^2 = (y - y_0)^2 + (x - l_1 \cos\theta)^2 \quad (2)$$

$$l_{xz}{}^2 = (z - z_0 - l_1 \sin\theta)^2 + (x - l_1 \cos\theta)^2 \quad (3)$$

$$l^2 = l_{xy}^2 + l_{xz}^2 \quad (4)$$

To verify the derived model, we commanded the end-effector to move in the form of geometric paths. We then recorded the movement using long-exposure snapshot which enables us to see the actual path taken. As a result, we found that the maximum available speed of the end-effector is 2 cm/sec. The speed beyond this number would result in linear actuator saturation and distortion in end-effector path.

3.6. Electronics

The control system for operation of robotic arms is for the front arm and the back arm. Each side consists of 3 linear actuators with DC motors and optical quadrature encoder to measure the position of the extended core. For the control system of motion, motors are grouped with 2 motors of each set. They are controlled via a Teensy 3.2 microcontroller which is a STM32 microcontroller with 96 MHz speed and with 2 quadrature encoder counters.

Operation and control board for the control system consists of Teensy 3.2 microcontroller and 2 motor driver boards VNH5019 assembled onto the circuit board, as shown in the Figure 15.

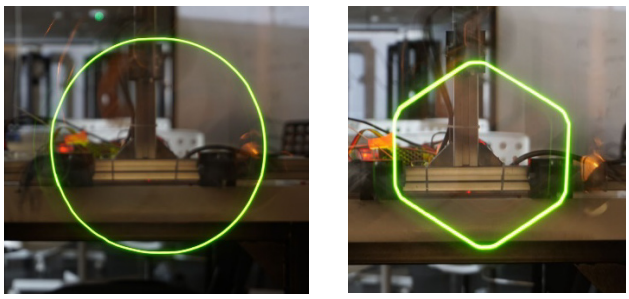


Figure 14: Actual path of the end-effector when tracing a 20mm radius circle at the speed of 2cm/s (left) and 4cm/s (right)

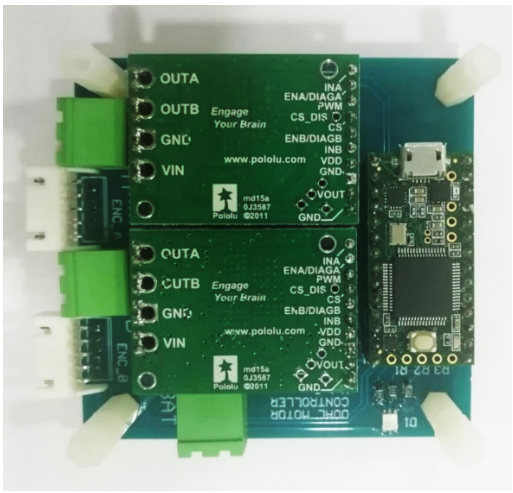


Figure 15: Operation and control board for the system

The control is commanded through the USB port to command the motor positon control program with PID control that is processed by microcontroller. The diagram for the operation of the motor control set is presented in the Figure 16.

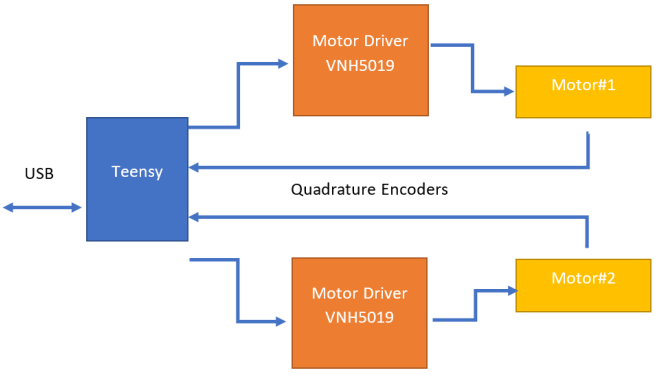


Figure 16: The diagram for the operation of the motor control set

All 3 motor control set are assembled altogether to control all 6 motors, and USB signal lines are connected through 4-port USB hub.

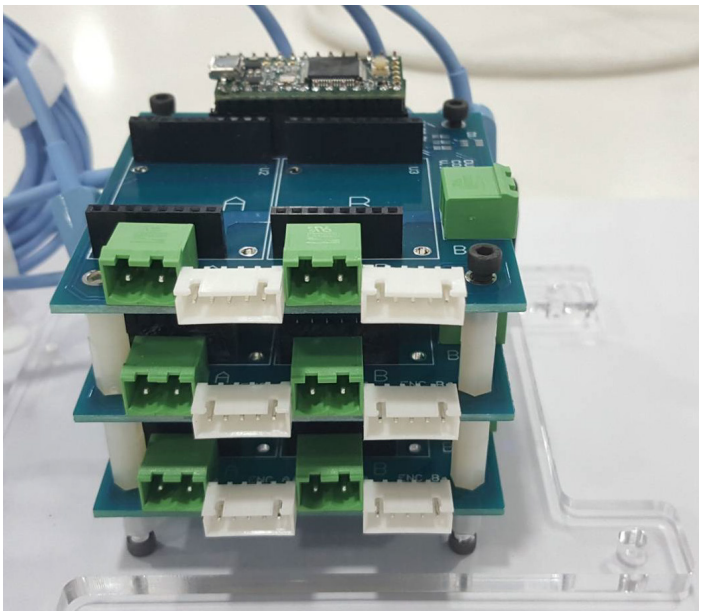


Figure 17: All motor control set are assembled altogether

Moreover, there is the fourth controller board. This board is used to control 3 BLDC motors. It is also equipped with inertial measurement module and load cells analog to digital converter to measure weight on each wheel for center of gravity calculation.

3.7. Software and control system

The main processor for this robot is an Industrial PC with Intel Core I5 processor running Ubuntu 18.04 as the operating system.

Operation and control program for the motion of the robot is constructed on the basis of Robotics Operating System (ROS) with the control structure in the Figure 18 below.

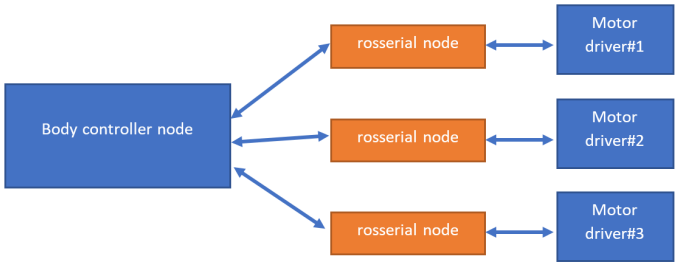


Figure 18: Operation and control program for the robot motion control



Figure 19: The developed inspection robot during set-up before the field test



Figure 20: Actual field test on a 115 kV transmission line

4. Results

The invented pilot-line pulling robot was utilized to line pulling operations at many 115-kV transmission lines for different of maintenance jobs during 2016-2019. The results showed its capability on pulling the lead line/conductor along the OHGW effectively and was accredited for real workings with no more harm to landowner and can operate with safety and reliability. Examined results after achieving the tasks have proved about the robot's performance on a pilot line pulling over the field with success. The robot can cut the personnel cost up to 34 less people and save the payment for damage's properties. For a direct output, the development of the robotic pilot-line pulling on the power line above the field can be achieved with no public concerns. For an indirect output, the electrical transmission network is safer and more dependable. Figure 7 demonstrates real implementation for inventive robot utilizations that pass the street, river, farmland, mountains, live switchyard, and electric power line. Also, the novel pilot line pulling robot was accredited as best practice by EGAT. The standards for invention, working, and maintenance were archived as references for transmission line operation and maintenance units nationally.

The inspection robot for transmission lines was specially designed for inspecting overhead power lines, running via the OHGW. Autonomous feature is a key challenge for developing the robot. Movement test was performed well on the simulated tower top for suspension type in the laboratory. After that, an actual field test was performed on July 2019 on suspension type and tension type towers of the real 115-kV power line. The observed results showed that the prototype of robot can run and transpose through actual obstacles like tested in the laboratory.

5. Conclusions

There is already evidence of different types of advanced robots designed and built by EGAT for the maintenance and inspection of high voltage transmission lines. The smart approach of work presently proposes the essential implementation of deploying innovative robots for approaches towards the smarter energy utility industry. The developed robots were successfully utilized to the actual transmission line in Thailand, causing tasks achieved with safety and reliability and without customer's service interruption. In addition, it helps improve effectiveness by saving manpower and be an effective approach rectifying inspection precision while comparing to the convention one, such as ground patrolling. By utilization with the smart robots, power companies will save on routine line maintenance and inspection

expenditure. Future development with highly advanced technology could be performed on the next generation of robot that can automatically climb the tower by itself.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The researches were supported by the Electricity Generating Authority of Thailand and the Center of Excellence in Mathematics, the Commission on Higher Education, Thailand.

References

[1] EGAT Northern Region Operation Division, “The optical ground wire inspector robot: phra-siva vehicle,” Conference of Electric Power Supply Industry, 2008.
[2] FIBO, “A prototype of inspection robot for water wall tubes in boiler,” EGAT Final Research Report, 2014.
[3] NECTEC, “Development of a robot with thickness less than 2 centimeters for generator inspection,” EGAT Final Research Report, 2014.
[4] MTEC, “Development of robotic system for welding,” EGAT Final Research Report, 2013.
[5] MTEC, “Automatic overlay welding robot system for power plant parts phase II,” EGAT Final Research Report, 2017.
[6] T. Disyadej, J. Promjan, K. Poochinapan, T. Mouktonglang, S. Grzybowski and P. Muneesawang, "High Voltage Power Line Maintenance & Inspection by Using Smart Robotics," 2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), Washington, DC, USA, 2019, pp. 1-4, doi: 10.1109/ISGT.2019.8791584.
[7] S. Montambault and N. Pouliot, "LineScout Technology: Development of an Inspection Robot Capable of Clearing Obstacles While Operating on a Live Line," ESMO 2006 - 2006 IEEE 11th International Conference on Transmission & Distribution Construction, Operation and Live-Line Maintenance, Albuquerque, NM, 2006, pp. , doi: 10.1109/TDCLLM.2006.340744.
[8] N. Pouliot and S. Montambault, "LineScout Technology: From inspection to robotic maintenance on live transmission power lines," 2009 IEEE International Conference on Robotics and Automation, Kobe, 2009, pp. 1034-1040, doi: 10.1109/ROBOT.2009.5152291.
[9] N. Pouliot, P. Latulippe and S. Montambault, "Reliable and intuitive teleoperation of LineScout: a mobile robot for live transmission line maintenance," 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO, 2009, pp. 1703-1710, doi: 10.1109/IROS.2009.5354819.
[10] S. Montambault, N. Pouliot, J. Toth and B. Spalteholz, "Reporting on a large ocean inlet crossing live transmission line inspection performed by linescout technology," 2010 IEEE International Conference on Robotics and Automation, Anchorage, AK, 2010, pp. 1102-1103, doi: 10.1109/ROBOT.2010.5509190.
[11] S. Montambault and N. Pouliot, "Field experience with LineScout Technology for live-line robotic inspection and maintenance of overhead transmission networks," 2010 1st International Conference on Applied Robotics for the Power Industry, Montreal, QC, 2010, pp. 1-2, doi: 10.1109/CARPI.2010.5624454.
[12] J. Toth, N. Pouliot and S. Montambault, "Field experiences using LineScout Technology on large BC transmission crossings," 2010 1st International Conference on Applied Robotics for the Power Industry, Montreal, QC, 2010, pp. 1-6, doi: 10.1109/CARPI.2010.5624413.
[13] N. Pouliot, D. Mussard and S. Montambault, "Localization and archiving of inspection data collected on power lines using LineScout Technology," 2012 2nd International Conference on Applied Robotics for the Power Industry (CARPI), Zurich, 2012, pp. 197-202, doi: 10.1109/CARPI.2012.6473341.
[14] S. Montambault, N. Pouliot and M. Lepage, "On the latest field deployments of LineScout Technology on live transmission networks," 2012 2nd International Conference on Applied Robotics for the Power Industry (CARPI), Zurich, 2012, pp. 126-127, doi: 10.1109/CARPI.2012.6473342.
[15] S. Montambault and N. Pouliot, "About the future of power line robotics," 2010 1st International Conference on Applied Robotics for the Power Industry, Montreal, QC, 2010, pp. 1-6, doi: 10.1109/CARPI.2010.5624466.
[16] P. Debenest et al., "Expliner - Robot for inspection of transmission lines," 2008 IEEE International Conference on Robotics and Automation, Pasadena, CA, 2008, pp. 3978-3984, doi: 10.1109/ROBOT.2008.4543822.
[17] P. Debenest and M. Guarnieri, "Expliner — From prototype towards a practical robot for inspection of high-voltage lines," 2010 1st International Conference on Applied Robotics for the Power Industry, Montreal, QC, 2010, pp. 1-6, doi: 10.1109/CARPI.2010.5624434.
[18] A. Phillips, E. Engdahl, D. McGuire, M. Major and G. Bartlett, "Autonomous overhead transmission line inspection robot (TI) development and demonstration," 2012 2nd International Conference on Applied Robotics for the Power Industry (CARPI), Zurich, 2012, pp. 94-95, doi: 10.1109/CARPI.2012.6473343.
[19] Tang Li, Fang Lijin and Wang Hongguang, "Obstacle-navigation control for a mobile robot suspended on overhead ground wires," ICARCV 2004 8th Control, Automation, Robotics and Vision Conference, 2004., Kunming, China, 2004, pp. 2082-2087 Vol. 3, doi: 10.1109/ICARCV.2004.1469485.
[20] Tang Li, Fang Lijin and Wang Hongguang, "Development of an inspection robot control system for 500KV extra-high voltage power transmission lines," SICE 2004 Annual Conference, Sapporo, 2004, pp. 1819-1824 vol. 2.
[21] Tang Li, Fu Shuangfei, Fang Lijin and Wang Hongguang, "Obstacle-navigation strategy of a wire-suspend robot for power transmission lines," 2004 IEEE International Conference on Robotics and Biomimetics, Shenyang, 2004, pp. 82-87, doi: 10.1109/ROBIO.2004.1521756.
[22] W. Hongguang, J. Yong, L. Aihua, F. Lijin and L. Lie, "Research of power transmission line maintenance robots in SIACAS," 2010 1st International Conference on Applied Robotics for the Power Industry, Montreal, QC, 2010, pp. 1-7, doi: 10.1109/CARPI.2010.5624428.
[23] X. Xiao, G. Wu and S. Li, "Dynamic Coupling Simulation of a Power Transmission Line Inspection Robot with its Flexible Moving Path when Overcoming Obstacles," 2007 IEEE International Conference on Automation Science and Engineering, Scottsdale, AZ, 2007, pp. 326-331, doi: 10.1109/COASE.2007.4341691.
[24] G. Wu, X. Xiao and Y. Lai, "A Wheel-Claw Hybrid Manipulator and its Grasping Stability for the Mobile Robot Rolling/Crawling Along Flexible Cable," 2007 International Workshop on Robotic and Sensors Environments, Ottawa, Ont., 2007, pp. 1-6, doi: 10.1109/ROSE.2007.4373964.
[25] J. Guo et al., "Database Environment of an Inspection Robot for Power Transmission Lines," 2009 Asia-Pacific Power and Energy Engineering Conference, Wuhan, 2009, pp. 1-4, doi: 10.1109/APPEEC.2009.4918249.
[26] G. Wu et al., "Design and Application of Inspection System in a Self-Governing Mobile Robot System for High Voltage Transmission Line Inspection," 2009 Asia-Pacific Power and Energy Engineering Conference, Wuhan, 2009, pp. 1-4, doi: 10.1109/APPEEC.2009.4918256.
[27] Gongping Wu, Tuo Zheng, Hua Xiao and Cheng Li, "Navigation, location and non-collision obstacles overcoming for high-voltage power transmission-line inspection robot," 2009 International Conference on Mechatronics and Automation, Changchun, 2009, pp. 2014-2020, doi: 10.1109/ICMA.2009.5245998.
[28] C. Hu, G. Wu, H. Cao and X. Xiao, "Obstacle Recognition and Localization Based on the Monocular Vision for Double Split Transmission Lines Inspection Robot," 2009 2nd International Congress on Image and Signal Processing, Tianjin, 2009, pp. 1-5, doi: 10.1109/CISP.2009.5303695.
[29] Gongping Wu, Tuo Zheng, Zhenglie Huang, Huan Liu and Cheng Li, "Navigation strategy for local autonomous obstacles-overcoming based on magnetic density detection for inspection robot of single split high voltage transmission line," 2010 8th World Congress on Intelligent Control and Automation, Jinan, 2010, pp. 6555-6561, doi: 10.1109/WCICA.2010.5554409.
[30] X. Xu, G. Wu, Y. He, X. Xiao, M. Liu and Q. Liu, "Navigation and control for a novel ground line inspection robot," 2010 IEEE International Conference on Mechatronics and Automation, Xi'an, 2010, pp. 1228-1232, doi: 10.1109/ICMA.2010.5589954.
[31] Naresuan University, “Development of mobile robot for inspection of high voltage transmission line,” EGAT Research Proposal, 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 500-511 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Alternative Braking Method for Small Scaled-Wind Turbines Connected DC Green House with Analogical Experiment on Blade Destruction

Anupa Koswatta[1], Faramarz Alsharif[2], Yasushi Shiroma[1], Mahdi Khosravy[3], Shiro Tamaki[1,*], Junji Tamura[2]

[1] *Graduate school of Engineering, University of the Ryukyus, 903-0213, Japan.*

[2] *School of Earth, Energy and Environmental Engineering, Kitami Institute of Technology, 090-8501, Japan.*

[3]*Media Integrated Communication Laboratory, Osaka University, 565-0871, Japan*

ARTICLE INFO

Article history:
Received: 01 February, 2020
Accepted: 25 February, 2020
Online: 04 April, 2020

Keywords:
Small-Scaled Wind Turbine
Wind turbine brake
Eddy current brake
Over-rotation
Strong wind

ABSTRACT

To supply a stable and efficient electricity, small-scaled wind turbine's brake system plays an immense role in various wind speeds. Small-wind power generation systems are difficult to operate in strong wind region since the turbine could be over-rotated and damaged if the brake system is not robust enough to maintain a stable angular velocity. Most of the small-scaled wind turbines use friction brakes to control the turbine speed for a stable electricity output. Since the friction brake are run out of time and needs frequent maintenance and eventually replacement, we introduce an eddy current based wind turbine brake system which is contactless with the rotor as an alternative to the friction brake system. The advantage of the proposed brake system is that the energy loss due to the friction will be reduced and will be more durable than the friction brake. The flow of this study is at first we did the analogical experiment of blade destruction to set to the maximum allowed angular velocity. Later, in order to verify the performance and stability requirement the mathematical implementation of eddy current brake system have done in DC-Green house for various wind penetration. Eventually the feasibility of eddy current brake system is confirmed in simulation results.

1. Introduction

The Wind energy provide a vast energy contribution to the modern world energy requirements. Recently its development made a huge impact to the world energy needs. Wind energy has a long history for more than 130 years [1]. In the early age of the wind industry, windmills were used to pump water for farms. Nowadays larger scale wind turbines can produce up to 10MW and Small wind turbines can produce less than 100 kilowatts. However, management of the wind energy sources has been always challenging as it has a considerable effect on different aspects of the power system such as the power quality [2], power harmonics and inter-harmonics [3], power line communications [4], etc. In this regard, research topics regarding how to obtain a maximum power output for high speed wind input without destroying the small-scaled wind turbine system can be considered as one of the important research areas of the wind energy. Because in high speed wind conditions the turbines output could be overloaded and could destroy the small-scaled wind turbine due to low inertia momentum. Therefore, in such wind condition, the wind operation must be stopped to avoid being damage to the small scaled wind turbine.

In this research, our work is about how to operate the wind turbine in such storm condition and how to generate the electricity without damaging the turbine. In the simulation results we have shown that it is possible to use eddy current magnetic brake system on the proposed method for the stall control. Wind turbine brake system plays an important role when there is storm condition. Most of the modern wind turbines use the friction brakes to control the rotation of the rotor of the wind turbine [5]. The brake pads are the important part to slower or stop the rotation of the rotor in strong wind situation. Therefore, these brake pads should be made for long term use. But still the modern wind turbine brakes can be used for few years, after that either it should be repaired or replaced by a new brake pad. Therefore, our aim is to introduce this magnetic brake system to replace the current friction brake system. The main specialty of this magnetic brake is there is no mechanical contact with the rotor for controlling the speed of the turbine. Therefore,

*Shiro Tamaki, Email: shiro@ie.u-ryukyu.ac.jp

this eddy current brake system is an alternative method to the conventional friction-based brake system [6]. Also, magnetic brake system required less maintenance than the conventional brake system [7].

In this paper, at first, we explain about small-scaled wind turbine blade which is destroyed due to experimental tension test that equivalent to centrifugal force due to over rotation of the blade. Afterwards the application of the magnetic brake and the eddy current brake were introduced. Later the mathematical modeling of the system and simulation results were explained.

2. Analogical experiment of blade destruction

Wind turbine blades are often destroyed by strong winds. The reason is due to the centrifugal force on the blades triggers into over-rotation by strong winds. In order to realize a tensile force equivalent to the centrifugal force until the blade destroys, the blade was subject to a tensile test. Figure 1 shows the test status of the broken wind turbine blade. In otherwords, this analogical experiment is done in order to obtain the safety margin of the blade. This experiment stands for loading force (i.e. equivalent to centrifugal force) to the blade until it breaks out. Through this experiment we can calculate the force that may cause destruction of the blade. Once the destructive force is obtained then we can re-calculate the maximum allowed angular velocity. The material of the blade is made by FRP (Fiber Reinforced Plastic). Figure 2 shows the experimental results of tension test of the blade. In this figure the horizontal axis stands for the displacement of the blade against the axis of the blade. Then vertical axis shows the tension force to the blade. Figure 2 was obtained using Microsoft Excel software. In this experiment, when the tensioned force reach to approximately 27 kN, at that point the blade is destroyed. So, in order to avoid such destruction, we have to concern safety margin. As it is obvious in Figure 2, tensioned force is linear with respect to displacement up to 6 mm. However, after 6mm the characteristics of the blade becomes non-linear against to force. Concerning safety margin and non-linearity of curve, we set the maximum allowed force to 4.5kN.



Figure 1. Broken wind turbine blade

Equation (1) shows the relation between tensioned force that is equivalent to centrifugal force and angular velocity as well the parameter and value are shown in table 1 below this equation. F_{max} stands for allowable maximum force, m is the weight and r is the length of blade.

$$F_{max} = mr\omega^2_{max} \quad (1)$$

Table 1: Specification of the blade

Parameter	Value	Unit
F_{max}	4.5	kN
m	4.5	kg
r	1	m
ω_{max} (without margin)	31.622	rad/s

According to equation (1) and table 1, the allowable maximum angular velocity ω_{max} is 31.622 rad/s (point A in the Figure 2). which is 4.5kN. We did not set the maximum angular velocity to the obtained value since safety margin is taken. Therefore, we put a secondary safety margin value as 29 rad/s (point B in the Figure 2) and primary safety margin value as 25 rad/s (point C). By doing so, we can obtain low penetration value 3.8kN for 29rad/s and 1kN for 25rad/s which is less than 4.5kN. From point A to point D can be considered as the blade destruction area. After point D the blade was broken. Hence, we set the ω_{max} as follows,

$$\omega_{max} = 25\text{rad/s}.$$

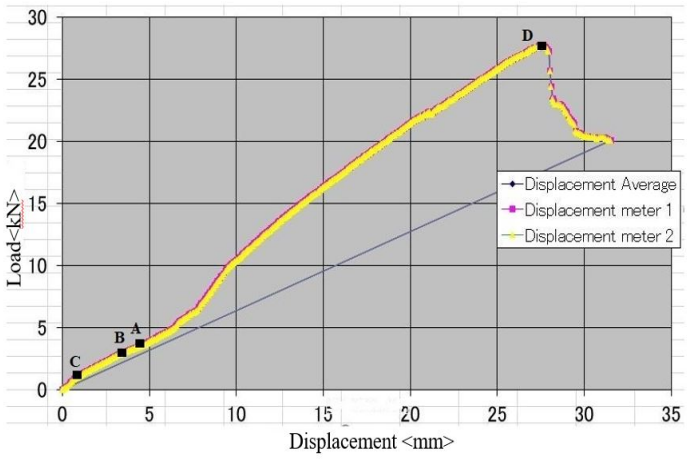


Figure 2. Blade destruction chart

3. Mechanism of eddy current brake

3.1. Eddy current theory

Some specific theories are related to the eddy current theory, such as the Faraday law of induction, the Lenz law and Fleming's right hand rule. Eddy current theory-based instruments are currently applied on many areas modern day. Some examples include bomb detectors and Electric meter for residence, among others.

Faraday law of induction states that a magnetic field can produce a current in a closed circuit when the magnetic flux linking to the circuit is changing [8]. Equation (2) introduces the mathematical expression of the Faraday law of induction,

$$\varepsilon = - \frac{d\Phi}{dt} \quad (2)$$

where, ε and Φ stands for electromotive force and magnetic flux, respectively.

Assuming this magnetic flux induced on a copper plate and its resistor is R_C and I is the induced eddy current on a copper plate. We can obtain the mathematical expression as (3), According to ohm's law,

$$-\varepsilon = R_c I_{eddy} \quad (3)$$

From (2) and (3) we can obtain the eddy current as the following equation indicates.

$$I_{eddy} = - \frac{1}{Rc} \varepsilon \quad (4)$$

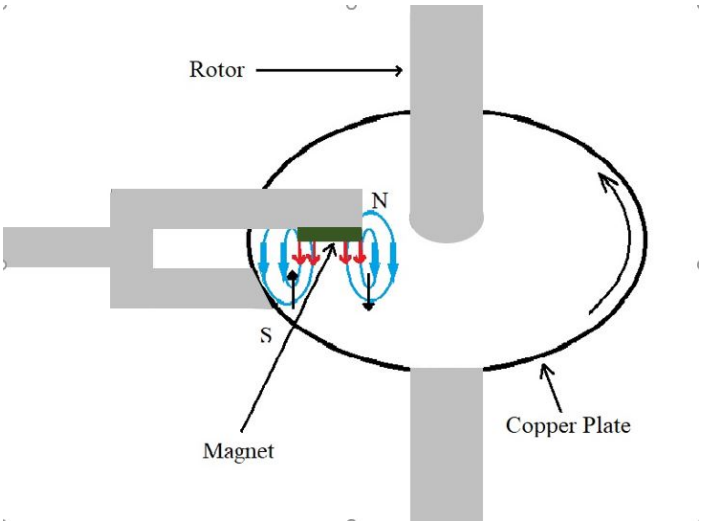
Therefore, it can be stated that induced eddy current can be calculated using (4).

3.2. Lenz law

Lenz law states that the direction of an induced electromagnetic force (e.m.f) is always such that it tends to set up a current opposing the motion or the change of flux responsible for inducing that e.m.f [9]. Eddy currents are the loops of currents which are induced due to the moving of a non-ferromagnetic conductor and a magnetic field according to the Faraday's law of induction as the equation above highlights. The path of the eddy current induced in the non-ferromagnetic material shaped in the form of a loop and it build up a magnetic field opposite to that of magnet [10]. The blue lines in the Figure 3 display the eddy currents induced on the conductor plate. The magnitude of the eddy current induced in the conductor is proportional to the size of the magnetic field. Also, the rotating copper plate which is attached to the rotor will exert a drag force from the permanent magnet. This drag force will operate as the brake force to control the angular velocity of the small scaled wind turbine. The reason for using the copper plate is that it's a non-ferromagnetic material. For ferromagnetic materials the eddy current will not be induced. Each time when the copper plate rotates the magnetic flux on the copper plate is changing. As a result of magnetic flux change the braking force will create due to the interaction of the eddy current on the plate and the magnetic field. Therefore, the magnet exerts a drag force on the copper plate. As a result of this, the rotation of the angular velocity of the copper plate decreases.

Nowadays electromagnetic brake systems are used widely such as in industrial applications, gym appliances, rollercoasters, recreation equipment's and maglev trains among others. This magnetic brake system is considered as a very efficient brake system because of its non-contact braking operation [11]. The advantage of this braking system is that, when the force become stronger, the magnetic force creates a stronger force against the system, relatively. Figure 3 shows how eddy current brake works. Only one section of the magnets was shown to describe the illustration more precisely. However, in the brake system incorporated multiple magnets to produce a strong braking force on the braking plate.

Figure 4 displays the breaking plate with induced eddy current. B1 and B2 are the induced magnetic field on the copper plate from the magnet.



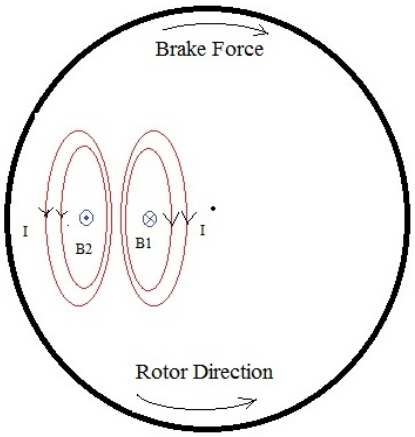


Figure 4. Braking plate

Mathematical expression for B is in (5). Here B denotes magnetic flux density and magnetic intensity denotes by H. μ represent the permeability of the atmosphere.

$$B = \mu H \quad (5)$$

The brake force is created as a result of the induced eddy current according to the Fleming's left-hand rule. While the rotor rotates counterclockwise the braking force will occur in clockwise direction.

4. Conventional method and proposed method and its application

Nowadays the conventional wind turbines use different types of control methods to get an efficient, stable and long-term durability power. Pitch control [12,13], yaw control, friction [14-16] control is popular control method are being used. In the friction control method, the brake pads are being used to control the angular velocity of the rotor. However, such brakes cannot be used long-term, since they need to be repaired or replaced if the brake pads were broken. Our proposed brake system, on the other hand, utilizes a magnetic eddy current brake system which controls the angular velocity in a non-contact manner with the rotor. In spite of the friction brake system that has the contact with the rotor, our

proposed method is contactless. Hence, it is adequate for long-term usage, rather than the time span offered by the conventional one.

The arrangement of magnets is displayed in Figure 5 in which the braking torque occurs in the opposite direction of the disk rotation.

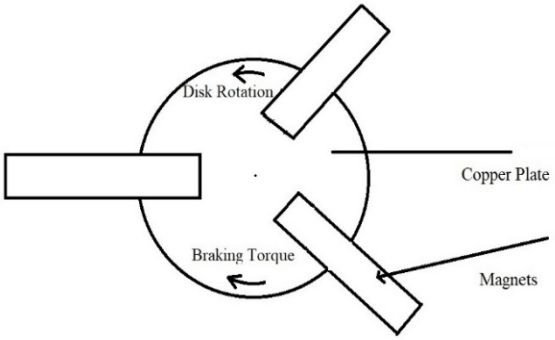


Figure 5. Cross section of the brake

When the small-scaled wind turbine angular velocity increases the copper plate angular velocity will increase simultaneously because the plate is attached to the rotor. When copper plate velocity increases the drag force on the plate from the induced eddy current also will be increased. Therefore, the angular velocity of the turbine will be controlled.

4.1. The structure of proposed brake system

In order to implement the eddy current contact less brake system, we have designed the following system shown in Figure 6. The principle is based on the eddy current occurrence that requires non-ferromagnetic low resistor material is implemented by the cooper plate and magnetic flux is set in the arm as it is shown in Figure 5. The principle of brake system is that when the angular velocity exceeds 25rad/s, the two magnets will be turned in between the copper plates in parallel using the arm that controlled by servo motor. As a result of this matter, due to rotation of plate and changing magnetics flux the emf is induced and eddy current will flow. Consequently, angular velocity of rotor will decrease due to the induced magnetic brake force. Hence, the excessive angular velocity will be restricted to the referenced value. Here, microcontroller has the controller role in the brake system. It measured the angular velocity and control the movement of arm by servo motor. By assumption of following prototype system, we construct the flow diagram of the system in next section.

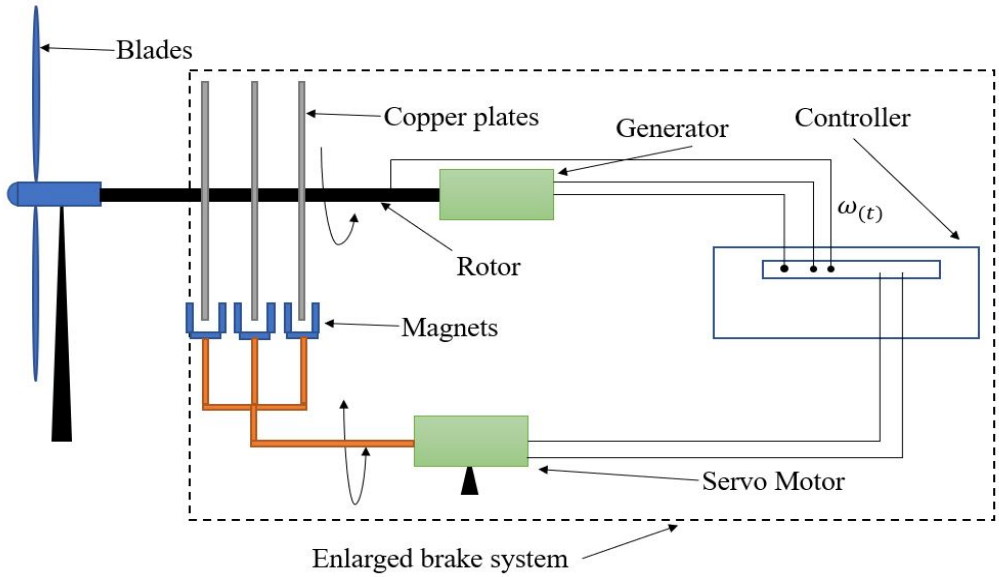
4.2. Flow diagram of the system

After the small-scaled wind turbine received the input wind, the flow will divide into mechanical power (P_{mech}) and power captured from wind (P_{ω}). Then the subtraction of power captured from wind and mechanical power can obtain electrical power (P_e). Afterward P_e will go through 3 φ power generator (P_g). Then the power received from the AC generator will distributes to AC/DC converter which includes rectifier and low-pass filter. At this stage the alternate current will convert to the direct current and it will charge the batteries for the purpose of supplying electricity to the DC green-house system. The load is variable. This process is displayed in the Figure 7.

When the measurement system receives the angular velocity $\omega_{(t)}$ of the small-scaled wind turbine the measurement system will send the measured angular velocity ω_m to the system controller. Afterwards the system controller will obtain the ω_{error} by calculating the difference between the reference value of angular velocity ω_{ref} and ω_m. Then, the system controller will turn on the servo motor to rotate the magnet closer to the braking plates. At this point, the magnetic brake force f_m will reduce the rotation of the brake plate as a result of the induced eddy current on the copper plate. This process was illustrated in Figure 8. In the next section mathematical modeling of entire system is described in detail.

4.3. Mathematical modeling

In this section the entire dynamics model of system is introduced. The system consists of small-scaled wind turbine, mechanical shaft, 3φ generator, DC converter and the DC bus in order. The DC bus is providing power to the loads at the end. The pattern of loads is shown in later section.



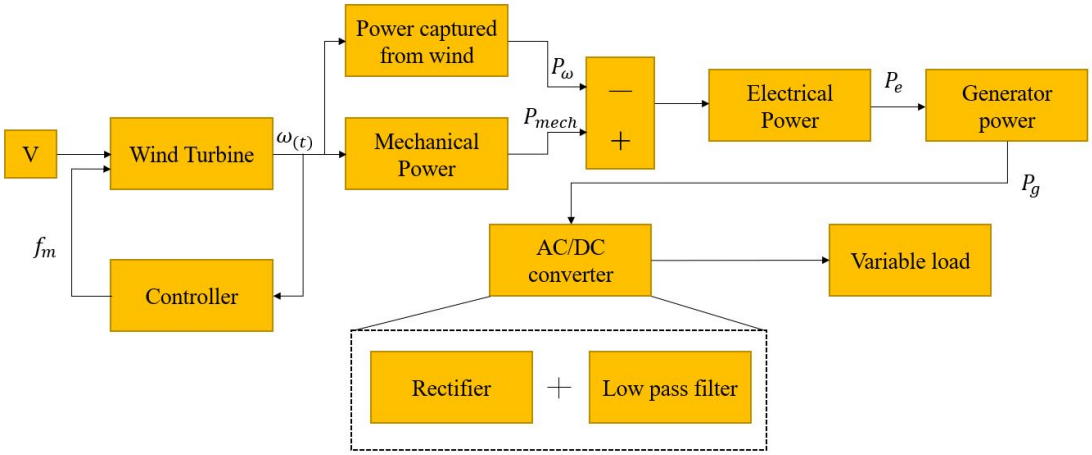


Figure 7. Flow diagram of the entire system

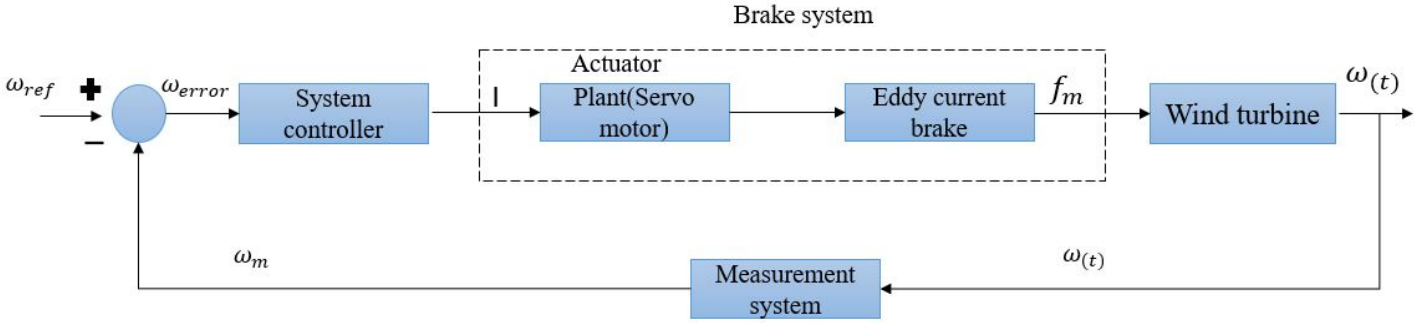


Table 2: Parameters

Symbol	Name
J_ω	Inertia moment of the blade
ω	Angular Velocity
ρ	Density of atmosphere
A	Area swept by blade
R	Length of blade
$V_{(t)}$	Speed of wind
f_r	Axis friction factor
μ	Permeability
q_m	Magnitude of the magnetic field of the magnet
q_c	Magnitude of the magnetic field of the copper plate
d	Distance between copper plate and magnet
a_{adj}	Decision functional adjustment
ω_{ref}	Reference value of angular velocity
ω_m	Measured angular velocity
I	Eddy Current
C_P	Power coefficient
C_t	Torque coefficient
η_g	Generator power coefficient
η_{DC}	Rectifier efficiency of DC converter

At first, we introduce the dynamics of the conventional model of small-scale wind turbine and afterwards the dynamics of the proposed method will be shown. The table 2 shows the parameters identification of the whole system. Additionally, the power efficiency of 3 φ generator and DC converter is considered as they are connected to the turbine shaft directly without coupling any gear.

The conventional dynamics of equation (6) represents the dynamics of the conventional small-scaled wind turbine system

$$J_\omega \frac{d\omega(t)}{dt} = \frac{1}{2} C_T \rho A R V_{(t)}^2 - f_r \omega_{(t)} \quad (6)$$

In order to show the stability of the system we apply the Laplace transform and made two assumptions to linearize the system

Assumption 1; $V_{(t)}$ is a constant value,

Assumption 2; $\omega_{(0)} = 0$.

Since velocity is constant the Laplace transform of the input term can be considered as $U_{(s)}$.

Here, s is the Laplace operator

The Laplace transform of angular velocity and its derivative can be obtain by following transformation rules

$$\mathscr{L}[\omega(t)] = \Omega_{(s)}$$

$$\mathscr{L}\left[\frac{d\omega(t)}{dt}\right] = s\Omega_{(s)} - \omega_{(0)} = s\Omega_{(s)} \quad .$$

Therefore, the Laplace transform of equation (6) is (7)

$$J_\omega s\Omega_{(S)} = U_{(S)} - f_r \Omega_{(S)} \quad (7)$$

Through equation (7) we obtain the transfer function of conventional method shown in (8).

$$\Omega_{(S)} = \frac{1}{J_\omega s + f_r} U_{(S)} \tag{8}$$

For (8) the bode plot is shown in Figure 9 to describe that how angular velocity responds to wind velocity that may contains various frequency components.

The parameter of transfer function is set to as following:

$$f_r = 0.2;\ J_\omega = 12;$$

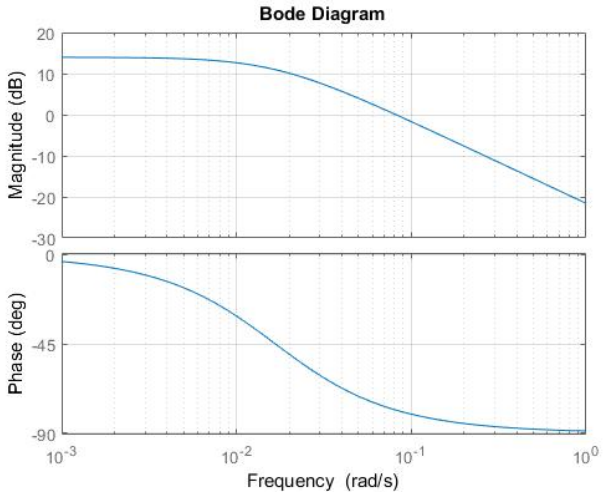


Figure 9. Bode diagram of transfer function of (8)

As it can be seen in Figure 9, the gain for low frequency components the output of transfer function will include up to 15 *dB*. This means that if there is a constant wind velocity (rated velocity) penetrates the turbine, the angular velocity increases rapidly. However, the natural wind velocity spectrum varies from low to high frequencies. As it is clear, in high frequency domain the gain is decreasing drastically. Therefore, the analysis of wind velocity spectrum [17] becomes significant that how affect the turbine rotation behavior. However, in this study we are not go through this.

Next for the stability analysis, from the transfer function we can obtain the pole of conventional method that the denominator of equation (8) is set to be 0. Hence, we have as following:

$$J_\omega s + f_r = 0$$

Eventually the pole of system (8) is

$$s = -\frac{f_r}{J_\omega} < 0 \quad .$$

The system is stable, since the pole is negative. So, we can conclude that the small-scaled wind turbine system has a stable behavior. The proposed method is shown by applying with combination of the eddy current and decision function.

The equation below expresses the proposed torque equation.

$$T_B = \frac{1}{2} C_T \rho A R V^2_{(t)} - T_{axis} - f_m f_{(\omega(t))} \tag{9}$$

The torque equation is shown in Figure 10.

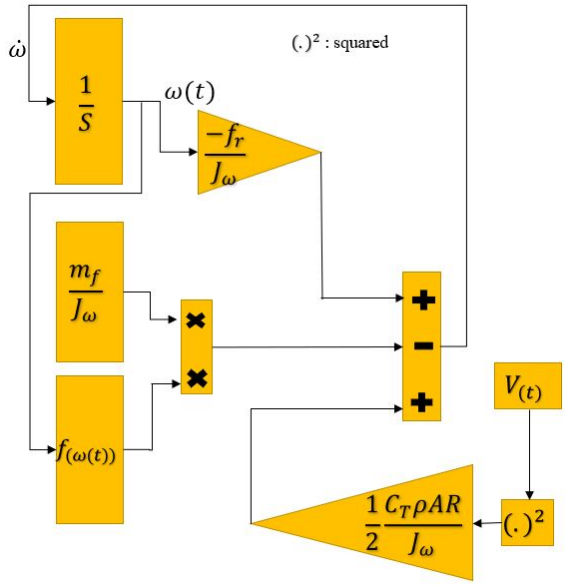


Figure 10. Flow of the torque equation

As it is clear in the above equation, conventional method of break system to restrict the angular velocity has done by T_{axis} which the technique was to varying the friction by brake pad. However, in this study since the $-f_m f_{(\omega(t))}$ term is proposed as contactless break method, the T_{axis} is considered as the friction of shaft coupled to the generator which can be assumed to be a constant value. Therefore, for the sake of contactless purpose the magnets are utilized in order to magnetize the copper plate and as results the eddy current will flow. The torque earned from term *fm* is single set permanent magnetic brake torque on the rotor which is

$$f_m = \mu \frac{2 q_m q_c}{4\pi d^2} \tag{10}$$

Where, *d* is the distance from copper to each magnet and q_m and q_c stands for magnitude of the magnetic field.

For the sake of stronger torque creation from magnetization, in this study we set 3 sets of permanent magnetic brake in parallel as described in Figure 5, the above equation becomes as follows

$$f_m = \mu \frac{6 q_m q_c}{4\pi d^2} \tag{11}$$

Furthermore, equation (12) represents the decision function [18-19].

$$f_{(\omega(t))} = \frac{1}{\pi} \tan^{-1}\left(\alpha_{adj}\left(\omega_{(t)} - \omega_{ref}\right)\right) + 1/2 \tag{12}$$

The decision function acts as a switch to brake system to decide at what moment the $\omega_{(t)}$ exceeds ω_{ref}. Therefore, whenever, angular velocity exceeds the reference value it will be activated. However, the performance of decision function depends on the coefficient α_{adj} which is analyzed in the later section.

Next , T_B is the torque of the blade which can be written as below equation.

$$T_B = J_\omega \frac{d\omega(t)}{dt} \tag{13}$$

Where, J_ω is the inertia moment of the blade.

Finally, T_{axis} represents the axis friction torque as described as follows

$$T_{axis} = f_r \omega_{(t)} \quad (14)$$

Where, f_r is axis friction factor.

It is notable that in conventional brake system this term was controlled to regulate and restrict the angular velocity [20]. However, we are not going through it in this paper.

The application of this study is to install 3 φ generator and rectifies its output to provide a DC power supply to the loads. Load considered to be DC-Green house that will describe the details in later section. The schematic of rectification and smoothing of its output, in other words the conventional DC converter consists of 6 diodes, capacitor and resistor which these two elements are used as the low-pass filter to smoothen the rectified output waves. Figure 11 represents a schematic of the generator which considered to be 3 φ power supply, full wave rectifier and low pass filter and a battery parallel with loads.

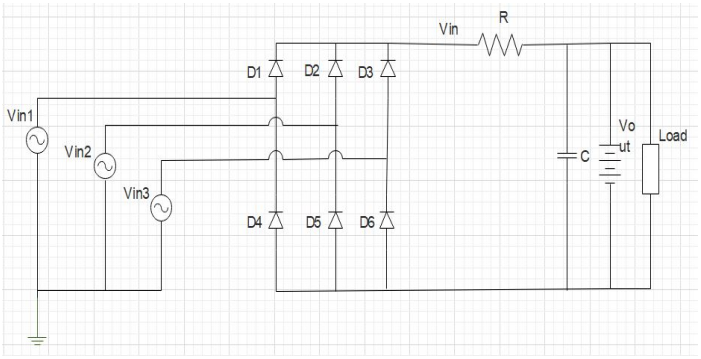


Figure 11 Three phase voltage circuit diagram with rectifier and low pass filter.

Equation (15) shows the low pass filter dynamics.

$$V_{in} - V_{out} = RC \frac{dV_{out}}{dt} \quad (15)$$

From (15) we can obtain the transfer function of low pass filter shown in following equation (16),

$$G_{LPF} = \frac{1}{1+RCs} \quad (16)$$

Where, s is Laplace operator.

As it can be seen it is significant to choose the R and C value to obtain a stable output to the loads. However, in this study, we are not going to the detail of designing the low-pass filter.

Now let us consider the power dynamics in order to obtain the anticipated generated electrical power from generator. According to the following equation the subtraction of P_{mech} by P_e is equivalent to $J_\omega \omega \frac{d\omega(t)}{dt}$. In other words, mathematical expression is as follows

$$J_\omega \omega \frac{d\omega(t)}{dt} = P_{mech} - P_e \quad (17)$$

Since P_{mech} is obtained by wind turbine characteristic formula, P_e can be anticipated by simple manipulation of above equation.

$$P_e = P_{mech} - J_\omega \omega \frac{d\omega(t)}{dt} \quad (18)$$

As mentioned previously, in this study since the efficiency is considered we define an output power efficiency of a 3φ AC generator and its rectification since the load is DC bus.

The output power of the 3φ AC generator displayed in (19).

$$P_G = \eta_g P_e \quad (19)$$

Here, η_g is the power coefficient of the generator and P_e is the electrical power received to the generator.

Further, equation (20) shows the converted DC power.

$$P_{DC} = \eta_{DC} P_G \quad (20)$$

where, η_{DC} represents the efficiency of rectifier of DC converter.

As considering the power stability analysis, it is significant that the produced power from wind turbine has adequate capacity of power in order to provide power to the loads. Therefore, the final stage of power conversion that is P_{DC}, this should satisfy the power to the loads constantly and globally in time domain which includes the transient and steady response.

Suppose that we have a function such as following (21)

$$\Gamma(P_{DC}, P_L) = \frac{P_{DC} - P_L}{P_{DC}} \quad (21)$$

which $\Gamma(.,.)$ stands for remnant function. The remnant function is composed of remnant power normalized by final stage power output that is DC converter output. Here, if and only if Γ is Semi Positive-definite function then the power supply is globally stable. The Below equation is used to verify the power supply stability.

$$\Gamma(P_{DC}, P_L) \geq 0 \quad (22)$$

Once the above equation is satisfied, then the required capacity is adequate for entire system. Besides, for above equation, in any situation $\sup_{\forall t} \Gamma(P_{DC}, P_L) \leq 1$ is satisfied so that the load's installation percentage can be determined. The satisfactory analysis and evaluation is done in later section as simulation.

4.4. Extension of the proposed method

From now on we will explain the extension of the proposed method which is about adding C_p and C_t as variables into the system. In equation (23) λ stands for tip speed ratio which is equal to the linear velocity of the wind turbine blade divided by the wind velocity. ω is the angular velocity of the blade and R is the radius of the blade [21].

$$\lambda = \frac{\omega R}{V} \quad (23)$$

Equation (24) represents the power coefficient of the wind turbine. Here, β is the blade pitch angle. We assume the blade pitch angle value is 0 for this study.

$$C_p = 0.5176 \left(\frac{116}{\lambda_i} - 0.4\beta - 5\right) e^{\frac{-21}{\lambda_i}} + 0.0068\lambda \quad (24)$$

Where $\frac{1}{\lambda_i} = \frac{1}{\lambda + 0.08\beta} - \frac{0.035}{\beta^3 + 1}$.

In this research pitch angle is assumed to be 0 constantly. Therefore,

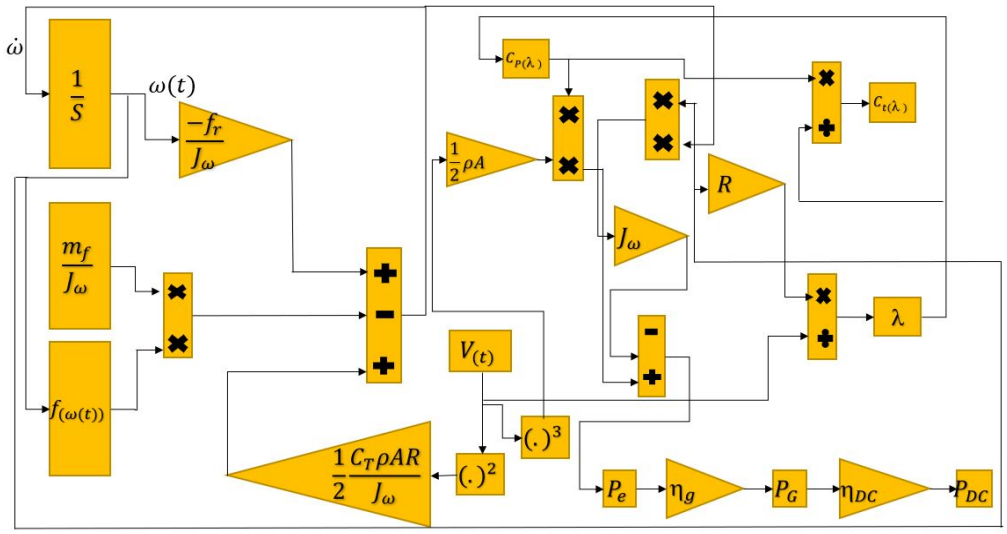


Figure 12. Expanded block diagram including C_p and C_t

$$\lambda_i = \frac{\lambda}{1\text{-}0.035\lambda}.$$

Equation (25) displays the torque coefficient

$$C_t = \frac{C_p}{\lambda} \quad (25)$$

The extended flow diagram of the [22] article is displayed in Figure 12. Note that $(.)^2$ and $(.)^3$ represent the mathematical operation as squared and cubed of input, respectively.

4.5. Analysis of decision function behavior

The decision function has an important role to do as regulation of angular velocity while it is operating. When angular velocity is greater than the reference value the decision function value will be greater than 0.5. Then the eddy current brake system will activate to reduce the angular velocity of the rotor.

For behavior analysis of decision function, let's give an artificial set of angular velocity that is ω = 2× t which t stands for time as shown in Figure 13. We used MATLAB/Simulink software to plot the following figures shows in this section.

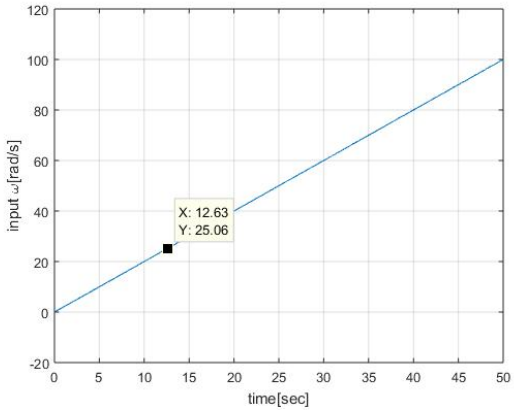


Figure 13. Input wind

For performance analysis of decision function a_{adj} is a critical coefficient. Therefore, we applied different a_{adj} values such as 0.1, 0.5, 1, 10, 200 to observe the output behavior of decision function. The triggered angular velocity value is set to be 25 rad/s. For these different a_{adj} values, the behavior of the decision function with respect to time is shown in Figure 14. For high values of a_{adj} such as 200 the decision function value is almost 1 at approximately 12 seconds which also can be seen in Figure 13 that at 12 sec angular velocity reaches 25 rad/s. According to Figure 15, when the a_{adj} is 0.1 the decision function's output has lower value comparing to the other a_{adj} values that are greater than unity. The decision function has a lower value in the beginning means that the brake system has already started to work when the angular velocity is below 25rad/s as shown in Figure 15. When the decision function output value is at 0.5 the brake system is triggered. When the a_{adj} increases the decision function value is converging to 1 for brake mode. That means the system must use a higher a_{adj} value in order to get an efficient brake control system. According to Figure 15 the highest a_{adj} value is 200. As it is clear in Figure 15, the area of non-braking mode except α_{adj} = 200 line (ideal mode) is shown in the gray color mesh and other area is for the braking mode.

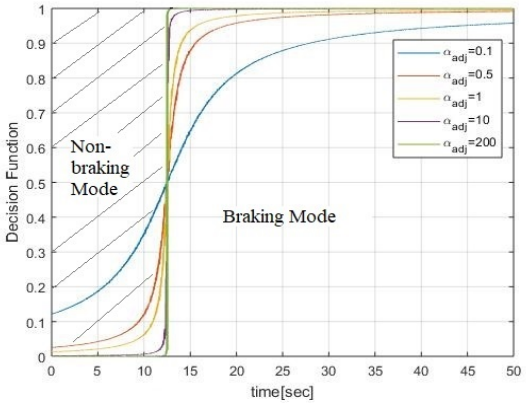


Figure 14. Performance of decision function against time

However, in reality it is hard to realize a system with a_{adj} 200 due to microcontroller process time and may not have enough capability to calculate speedily. Therefore, some internal delay

may occur in the system and stability and performance analysis should be done for system including time delay element [23-25]. However, in this study, since we focus on behavior of brake system performance, instead of analyzing the delay system we set the a_{adj} to an acceptable value. Hence, we applied the a_{adj} as 10 to which has adequate transient response according to Figure 14 and at the same time is considered to be a reasonable value for implementation in actual case.

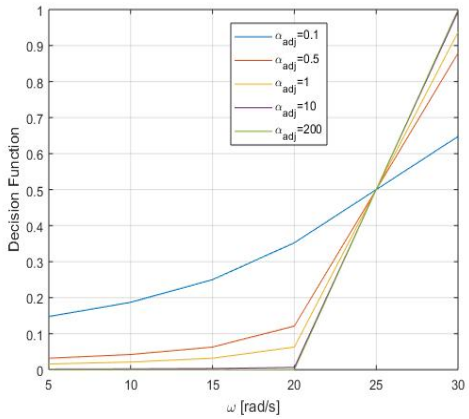


Figure 15. Decision function against angular velocity

5. Simulation and Results

In this section, first, we built the mathematical model of the control system then we simulated it in the MATLAB/Simulink based environment. At first we simulate the angular velocity response for different mean velocities from 10 to 30 m/s. Next we set two patterns of input wind velocity and observe the angular velocity behavior. One set is a regular rising and declining wind velocity (pattern No.1) and the other one is a random one (pattern No.2) which are shown in results. Additionally, for each pattern, in order to observe the decision function performance, we have shown the offsets between the ω_{ref} and ω_{steady} with respect to α_{adj}= 6 to 200. The offset stands for the subtraction of reference angular velocity by steady response which concerns system performances. Eventually, for both patterns, to observe the power provision stability evaluation function $\Gamma(P_{DC}, P_L)$ is shown with regard to time response.

5.1. Simulation conditions

In this simulation, power coefficient and torque coefficients are considered as variables. Previously in [1] power coefficient and torque coefficient were constants. Simulation time for this research is 1000s. Sampling time is 0.1s. The load as shown in Figure 16 is applied to entire simulation. The input load behavior changes depending on various conditions in DC-green house such as weather, temperature, light condition. In the greenhouse the power consumption is varying through day and night time. Specially in night time the power consumption would be higher since many devices are in operation. Such as the LED lights which exposes light to products, microcontrollers and sensors. Therefore, we created the load that applies in real case to verify the proposed system that while restriction of angular velocity weather it is possible to provide an adequate power to the loads or not. However, for the sake of simplicity for analysis as it is expressed duration time is set to 1000 s.

Table 3 displays the parameters values with the SI units

Table 3: Values of the Parameters

Symbol	Value	Unit
J_ω	12	kg.m^2
ω	Variable	$rads^{-1}$
ρ	1.225	kg.m^3
A	3.14	m^2
R	1.5	m
$V_{(t)}$	Variable	ms^{-1}
f_r	0.2	kg. $m^2 s^{-1}$
μ	$4\pi10^{-7}$	T. m. A^{-1}
q_m	$62.5*10^9$	A.m
q_c	$0.1*10^{-4}$	A.m
d	0.05	m
a_{adj}	10	
ω_{ref}	25	$rads^{-1}$
η_g	0.9	
η_{DC}	0.9	

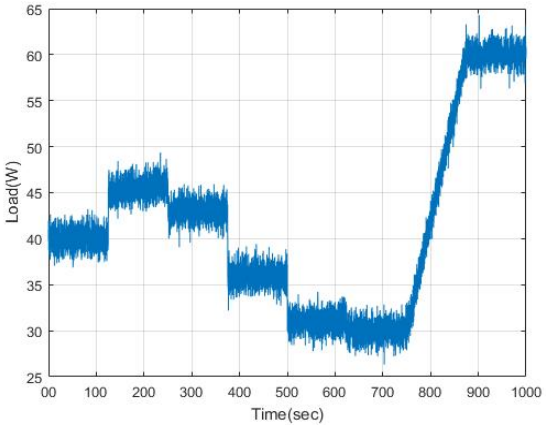


Figure 16. Pattern of the load

5.2. Results

From now on, Figure 17 to Figure 21 displays the angular velocity behavior for different input average velocities of 10, 15, 20, 25 and 30 m/s in order with controller decision function coefficient is set to α_{adj} = *10* and the stall angular velocity or the referenced one is considered to be 25 rad/s as it has been chosen in section 2.

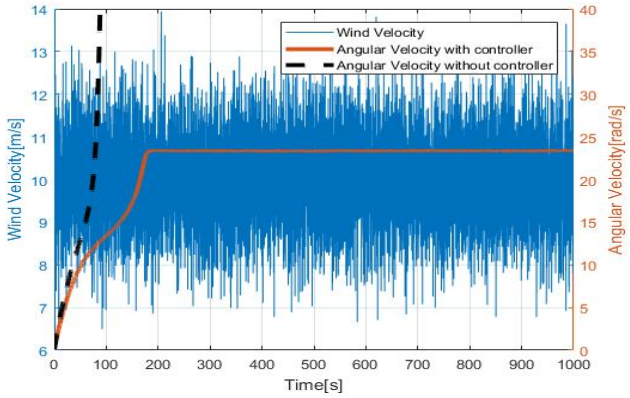


Figure 17. Stall control chart for average wind 10m/s

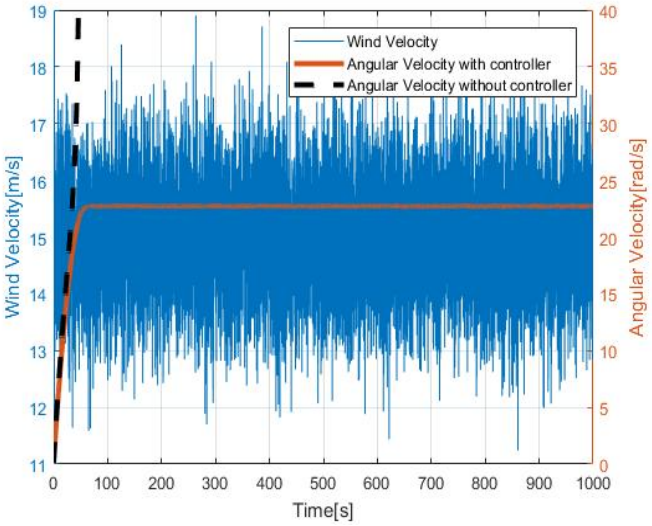


Figure 18. Stall control chart for average wind 15m/s

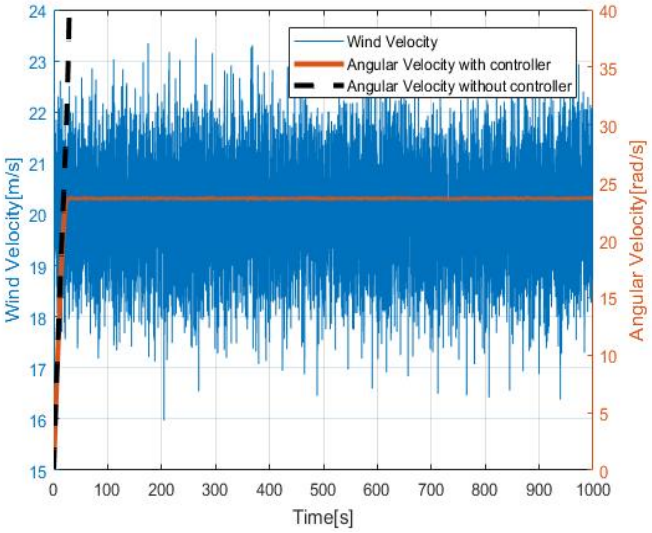


Figure 19. Stall control chart for average wind 20m/s

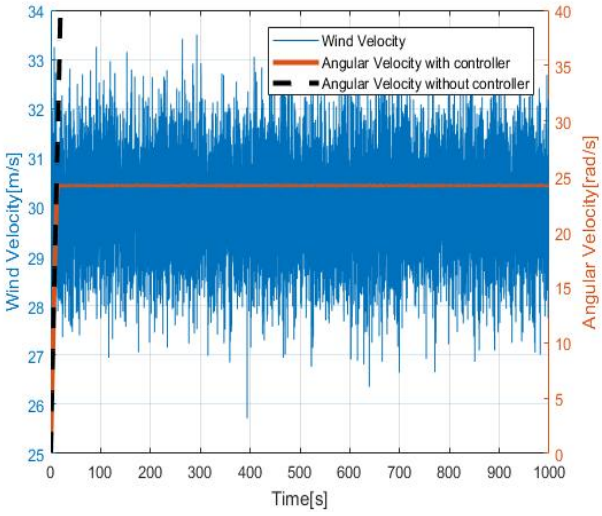


Figure 20. Stall control chart for average wind 25m/s

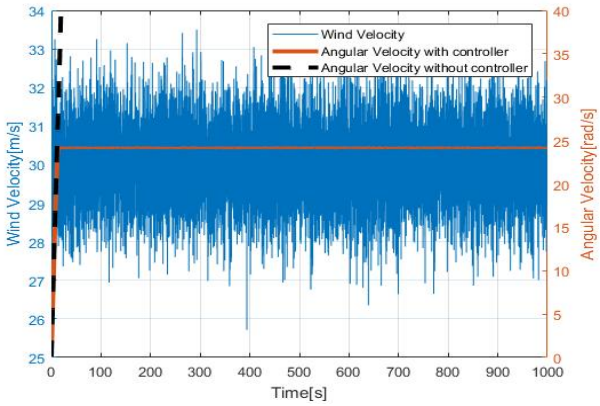


Figure 21. Stall control chart for average wind 30 m/s

By analyzing the Figure 17 to Figure 21 it is obvious that the magnetic brake system restricts the angular velocity up to 25 rad/s for average wind velocities of 10m/s to 30m/s. However, without controlling it will diverge and system becomes unstable. Moreover, the transient response of decision function seems to be adequate. Therefore, we can conclude that $\alpha_{adj} = 10$ has a stable behavior and acceptable performance. Following Figure 22 summarize the angular velocity behavior when the input wind velocity changes from 10m/s to 30m/s.

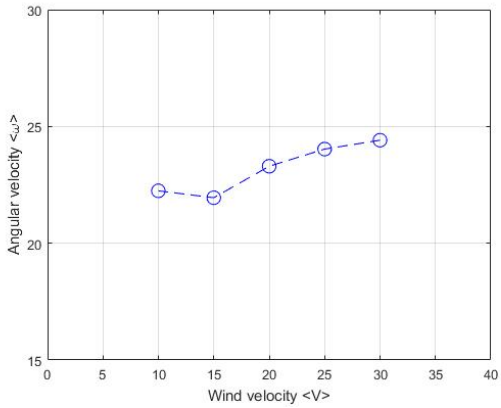


Figure 22. Wind velocity from 10m/s to 30m/s against angular velocity

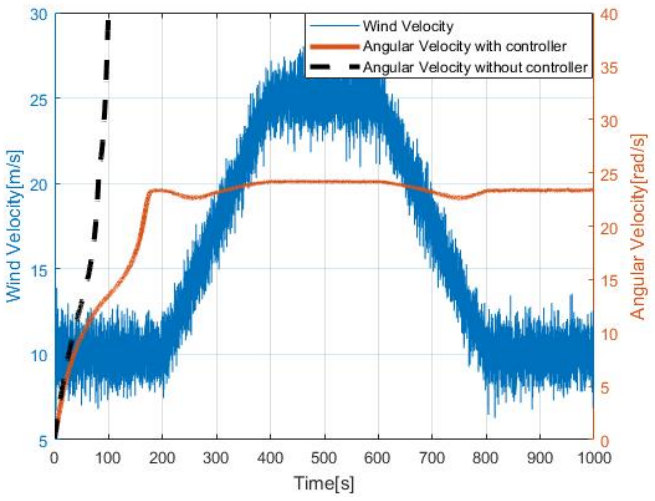


Figure 23. Angular velocity behavior for wind velocity pattern No. 1

Figure 23 and 24 explain how the brake system works with and without the proposed control. Angular velocity without the brake control is displayed using the broken line which is diverging on the Y-axis for both wind patterns. The maximum average wind velocity in this simulation is 25m/s and the minimum value is 5m/s. Patterns of input velocity have been created artificially.

As it is shown in above figure, the behavior of angular velocity with proposed controller against the input wind velocity pattern No.1 has a stable attitude with respect to time. As well, it is visible that the angular velocity is controlled up to 25rad/s even during high wind penetration conditions which is during 400 to 600 sec.

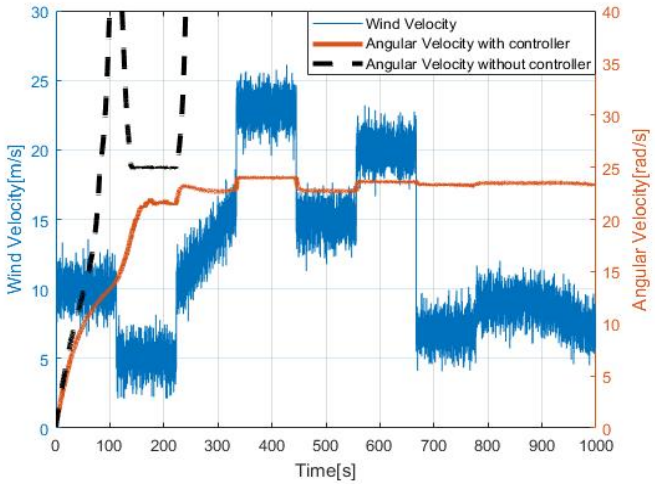


Figure 24. Angular velocity behavior for wind velocity pattern No. 2

In Figure 24, it is clear that even though angular velocity with proposed controller fluctuates but it has stable response and have been suppressed up to 25 rad/s and cease the excessive rotation. The arise of fluctuation of angular velocity is due to the wind pattern that have been penetrated to the wind turbine. Therefore, the angular velocity behavior should be enhanced for such sudden changes of wind velocity. However, for now it satisfies the requirement. Therefore, mechanical point of view, we can conclude that the angular velocity can be sustained to a desired or referenced value by proposed method.

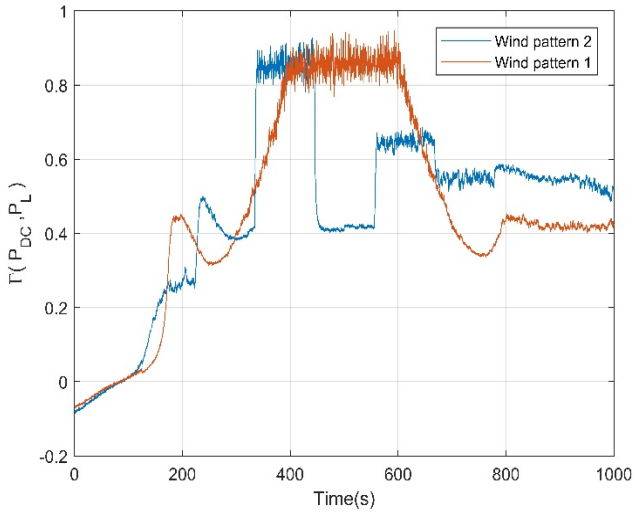


Figure 25. Time response of evaluation function for wind pattern 1 and 2

Next the power provision is evaluated by evaluation function which has been defined previously. Following Figure 25 is shown the evaluation function for the wind velocity pattern No. 1 and 2, respectively.

As it is obvious, except the transient duration, equation (21) is satisfied. Hence, the proposed controller can also provide an adequate power supply to loads and restricts the angular velocity. Also, the power output contains the remnant quantity of power particularly during high wind penetration that means it can be stored to power bank or charges the peripheral batteries.

Finally, in order to ensure the decision function performance that is part of controller the offsets of referenced angular velocity and steady value of angular velocity with controller has been plotted with respect to α_{adj} in Figure 26.

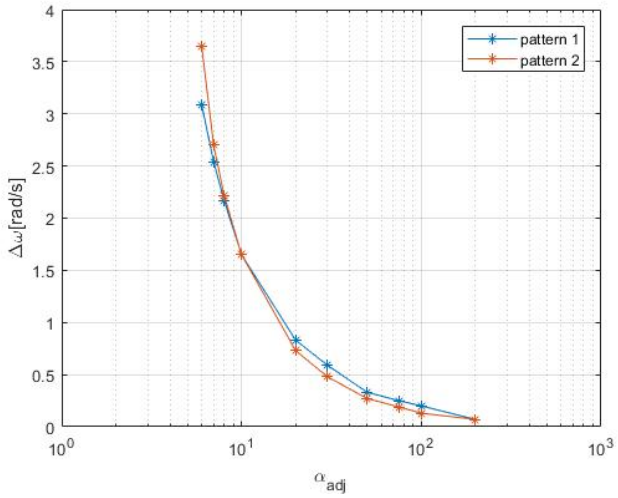


Figure 26. Offset for average wind velocity 20m/s

As it is clear in above figure, if the $\alpha_{adj} < 10$, then the corresponding offset increase. On the other side, if $\alpha_{adj} > 10$, then the corresponding offset would decrease. This means for higher value of factor, performance of the system improves since the offset decreases. However, as have been discussing, it is fairly hard to set α_{adj} to high value due to microcontroller processor capability for realization. Therefore, α_{adj} can be set to *10* according to the above since it is not too high value and at the same time it has sufficient performances.

6. Conclusion

In this paper, we have simulated an eddy current magnetic brake system for the small-scaled wind turbines. In strong wind condition, several types of control methods utilized to get the desired electricity output for small-scaled wind turbines. But still the modern wind turbines unable to get a stable generated electricity output when there is a storm condition. In storm condition, most of the wind turbines stop their operation to protect the turbine and it's all components. As we have seen from the simulation results it is possible to conclude that using the eddy current brake system, we can limit the angular velocity below 25rad/s when there is storm condition and satisfies the power provision to the loads, simultaneously. Since there is no mechanical connection with the rotor and the brake system, we

believe it may be able to use for much longer time than the conventional friction brake system. Thus, for future works, we will apply the Lyapunov function for the stability analysis. Moreover, in this paper the trigger of angular velocity is set based on physical destructive value. However, maximum power point should also be considered for enhancing the efficiency. Therefore, we will use the maximum Power Point Tracking Method (MPPT) to get a maximum power output form objected wind turbine. Additionally, pitch controlling method is a very useful method to control the wind turbine operation. Since we haven't applied this method to the proposed system in this paper, we will apply pitch control method as future work for this system. Finally, the integration of pitch angle control and proposed eddy current brake system will be implemented.

Conflict of Interest

The authors declare no conflict of interest.

Reference

[1] J.F Manwell, J.G. Mcgowan, A.L. Rogers, Wind Energy Explained, John Wiley and Sons, 2009.

[2] E. Santos, M. Khosravy, M.A. Lima, A.S. Cerqueira, CA Duque, "ESPRIT associated with filter bank for power-line harmonics, sub-harmonics and inter-harmonics parameters estimation". INT. J. ELEC. POWER, **118**, p.105731, 2020. https://doi.org/10.1016/j.ijepes.2019.105731

[3] E. Santos, M. Khosravy, M.A. Lima, A.S. Cerqueira, C.A Duque, and A. Yona, "High accuracy power quality evaluation under a colored noisy condition by filter bank ESPRIT" Electronics, **8**(11), p.1259, 2019. https://doi.org/10.3390/electronics8111259

[4] A.A. Picorone, T.R. de Oliveira, R. Sampaio-Neto, M. Khosravy, M.V. Ribeiro, "Channel characterization of low voltage electric power distribution networks for PLC applications based on measurement campaign", INT. J. ELEC. POWER., **116**, p.105554. 2020. https://doi.org/10.1016/j.ijepes.2019.105554

[5] Y.B. Lee, G.C. Lee, J.D. Yang, J.W. Park, D.C. Baek, "Failure analysis of a hydraulic power system in the wind turbine" ENG. FAIL. ANAL. **107**, 104218, 2020. https://doi.org/10.1016/j.engfailanal.2019.104218

[6] R. Yazdanpanah, M. Mirsalim., "Hybrid electromagnetic brakes: design and performance evaluation" IEEE T. ENERGY CONVER., **30**(1), 60-69, 2015. https://doi.org/10.1109/TEC.2014.2333777

[7] H. Sodano, J.S. Bae, "Eddy current damping in structures" The Shock and Vibration Digest., **36**(6), 469-478, 2004. https://doi.org/10.1177/0583102404048517

[8] J. D. Kraus, K. R. Carver, Electromagnetics, 2nd ed. McGraw Hill, 1973.

[9] John Bird, Electrical circuit theory and technology. Third edition, Newnes, 2007

[10] M. O. Gulbahce, D.A. Kocabas, I Habir, "Finite element analysis of a small power eddy current brake" Proceedings of 15th International Conference MECHATRONIKA, Prague, Czech Republic, 2013.

[11] K.J. Lee, K. Park, "Optimal robust control of a contactless brake system using an eddy current" Mechatronics, **9**(6), 615–631, 1999. https://doi.org/10.1016/S0957-4158(99)00008-2

[12] F. Asharif, S. Tamaki, T. Nagado, T. Nagata M.R. Asharif, "Analysis of non-linear adaptive friction and pitch angle control of small-scaled wind turbine system," Future Generation Information Technology FGIT 2011, Control and Automation CA, Springer-Verlag, Lecture Note: Communication in Computer and Information Science, 26-35, December 2011. https://doi.org/10.1007/978-3-642-26010-0_4

[13] F. Alsharif, J. Tamura, S. Tamaki, K. Anupa, M. Futami, "Dynamics modification of passive pitch control for small-scaled wind turbine" Grand Renewable Energy 2018, Yokohama, Japan, June 2018. https://doi.org/10.24752/gre.1.0_147

[14] F. Asharif, S. Tamaki, H. Teppei, T. Nagado, T. Nagata, "Feasibility confirmation of angular velocity stall control for small-scaled wind turbine system by phase plane method" IEEK Transactions on Smart Processing and Computing, **2**(4), 240-247, 2013.

[15] F. Asharif, S. Tamaki, T. Nagado, M. R. Alsharif, "Comparison and evaluation of restrain control in wind turbine with various shock absorber considering the time-delay" International Journal of Control and Automation, Vol. **5**(3), 111-131, September 2012.

[16] F. Asharif, M. Futami, S. Tamaki, T. Nagado, K. Asato, "Stability and performance analysis of passive and active stall control of small-scaled wind turbine system by phase plane method" Proceedings of IEEE, International Conference on Intelligent Information and BioMedical Sciences (ICIIBMS 2015), .27-33, OIST, Okinawa, Japan, November 2015. https://doi.org/10.1109/ICIIBMS.2015.7439483

[17] E. Branlard, A. T. Pedersen, J. Mann, N. Angelou, A. Fischer, T. Mikkelsen, M. Harris, C. Slinger, B. F. Montes, "Retrieving wind statistics from average spectrum of continuous-wave lidar" Atmos. Meas. Tech., **6**, 1673–1683, 2013. https://doi.org/10.5194/amt-6-1673-2013

[18] F. Asharif, S. Tamaki, T. Hirata, T. Nagado, T. Nagata, "Design of electromagnetic and centrifugal force pitch angle stall controller on application to small-scaled wind turbine system" International Symposium on Stochastic Systems Theory and Its Applications, 2013. https://doi.org/10.5687/sss.2014.231

[19] F. Asharif, M. Futami, S. Tamaki, T. Nagado, K. Asato, "Stability and performances analysis of electromagnetic stall control system by phase plane method" 47th ISCIE International Symposium on Stochastic Systems Theory and Its Applications, Hawaii, USA, December 2015. https://doi.org/10.5687/sss.2016.345

[20] F. Asharif, S. Tamaki, T. Nagado, T. Nagata, M. R. Asharif, "Design of adaptive friction control of small-scaled wind turbine system considering the distant observation" International Conference of Control and Automation, LNCS Springer, 213-221, November 2012. https://doi.org/10.1007/978-3-642-35264-5_30

[21] N. Karakasis, A. Mesemanolis, C. Mademlis, "Wind turbine simulator for laboratory testing of a wind energy conversion drive train" 8th Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MEDPOWER), 2012. https://doi.org/10.1049/cp.2012.2033

[22] K. Anupa, F. Asharif, S. Yasushi, K. Ken, S. Tamaki, J. Tamura, "Contactless magnetic braking control unit for small-scaled wind turbines for dc green house" The 34th International Technical Conference on Circuits/Systems, Computers and Communications, Jeju Shinhwa World, Republic of Korea, 2019. https://doi.org/10.1109/ITC-CSCC.2019.8793443

[23] F. Asharif, S. Tamaki, T. Nagado, T. Nagata M. R. Asharif, "Application of internal model controller for wind turbine system considering time-delay element" 8th International Conference on Informatics in Control, Automation and Robotics, 199-202, Noordwijkerhout, Netherlands, July 28-31, 2011.

[24] F. Asharif, S. Tamaki, T. Nagado, T. Nagata, "Analysis of hybrid robust controller considering disturbance, noise and time-delay" 18th Iranian Conference on Electrical Engineering Isfahan, Iran, Proceeding of IEEE Conference 978-1-4244-6760-0/10, 11-13, May 2010. https://doi.org/10.1109/IRANIANCEE.2010.5506990

[25] F. Asharif, S. Tamaki, T. Nagado, T. Nagata, M. Rashid, M. R. Asharif, "Feedback control of linear-quadratic-integration including time-delay system" The 24th International Conference on Circuit-Systems, Computers and Communication, ITC-CSCC'09, 1054-1057, Jeju, Korea, July 5-8, 2009.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 4, 517-521 (2019)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Mechanical Behaviors of Kaolin Powder Filler Polypropylene/Low Density Polyethylene Blends

Pham Thi Hong Nga*

Hochiminh City University of Technology and Education, Mechanical Engineering Faculty, Metal and Technology Department, 70000, Vietnam

ARTICLE INFO

Article history:
Received: 23 June, 2019
Accepted: 14 August, 2019
Online: 23 August, 2019

Keywords:
Kaolin
PP/LDPE blend
Tensile strength

ABSTRACT

In this work, kaolin powder was filled in PP/LDPE blend as the filler with the amount from 0 to 14 wt%. The ratio of PP/LDPE was fixed as 50/50 in all the experiments. The tensile strength, impact strength, and hardness were investigated in according with ASTM. The results showed that the tensile strength of PP/LDPE blend was slightly increased, the hardness was also increased while the impact strength was decreased in the presence of kaolin powder as the filler from 0 to 14 wt.%.

1. Introduction

Mixing of two or more different polymers is now considered as an economical way to the development of new polymers [1]. Low-density polyethylene (LDPE) has good mechanical properties, withstand high temperatures, easy to handle so they have wide applications in the industry [2]. PP also has some characteristics such as high stiffness, low plasticity, but these characteristics making the PP structure are easily destroyed, so the PP applications are also limited [3]. Therefore, to improve the mechanical properties of PP, and to create new materials with appropriate characteristics, LDPE has been studied and combined with PP to form PP/LDPE blend composition. Although, PP is similar to LDPE, they are different in some significant properties, PP/LDPE blends produce immiscible form.

The recent work of L.F. Kadhim et al [4], who investigated the mixures of PP and LDPE in the following percentages of LDPE by weight: 25, 50, 75. The results showing that the addition of LDPE to PP, have been declined the tensile strength, flexural strength, flexural modulus and hardness while the density improved LDPE as a result of the nature of LDPE is more flexible than PP. However, they are still only used for business purposes. One of the most important ways of polymer mixing is the incorporation of fillers to enhance mechanical toughness. As fillers, kaolin, $CaCO_3$, and talc are used along with engineering polymers to reduce both the production costs and to improve the properties. PP/LDPE/filler blends have been studied by different researchers from different aspects [5-7]. In types of filler, kaolin has certain advantages to improve the characteristics of PP/LDPE blend composition. It changes bonding between polymer blends to enhance bonding between blends and to create a chemical bond course between the blends of polymers and kaolin, therefore to improve the mechanical properties of PP/LDPE blends [8-11]. Many studies have so far been done on investigating PP/LDPE/Kaolin composites. According to S.N. Mustafa et al, the addition of kaolin powder to the PP/LDPE blend leads to increase the tensile strength, modulus of elasticity, shore-D hardness and impact strength and it decreases the elongation at break [12]. Water absorption of the PP/LDPE/kaolin composites behaves as function of time has also been investigated, and it increases by increasing immersion time for the same filler content, while the absorbed amount of water increases, by increasing the wt.% of kaolin at constant immersion time.

Table 1: Compositions of the samples (wt.%)

Material	Components (wt.%)				
	Samples				
	S1	S2	S3	S4	S5
50% PP and 50% PE	100	97	95	90	86
Kaolin	0	3	5	10	14

2. Experimental

2.1. Materials

PP (LyondellBasell - Moplen HP500N, origin Saudi Arabia) and LDPE (SABIC - LDPE 4024, origin Saudi Arabia) supplied

* Pham Thi Hong Nga, Email: hongnga@hcmute.edu.vn

by Thuan Thang Plastics Co., Ltd. Kaolin ($Al_2Si_2O_5(OH)_4$) were collected from Tran Tien Chemical Company. Kaolin has a particle size of 44 µm, specific gravity 2.58 - 2.63g/cm^3, and whiteness 89 - 93%.



Figure 1: The sample after pressing

In the preparation of the blends, samples including PP/LDPE and kaolin were put into the mixer according to the ratios (Table 1). These composites are called as S1, S2, S3, S4 and S5. All compounds were prepared by using system of mixing and extruding machine Polylab OS - Haake (Germany). Basic parameters of the system are 20 cm^3 sealed mixing chamber, the device extracts two screws L/D = 25, D = 16mm, one-screw extruder connected with sheet extrusion system (0.2 - 1.2 mm x 100mm). Samples were mixed for 6~7 minutes at a mixing temperature of 180ºC. After mixing, the sample are pressed to plates with the temperature of the tray at 180ºC. The pressing process was carried out for 5 minutes and then cooled for 20 minutes. All of these blends were prepared as samples keeping the PP/LDPE (50/50) ratio constant. The sample after pressing is 50 x145 mm and the thickness is 3.2 mm (Figure 1). The samples are then cut to the size of the tensile test specimen and the impact strength test specimen.

2.2. Experimental methods

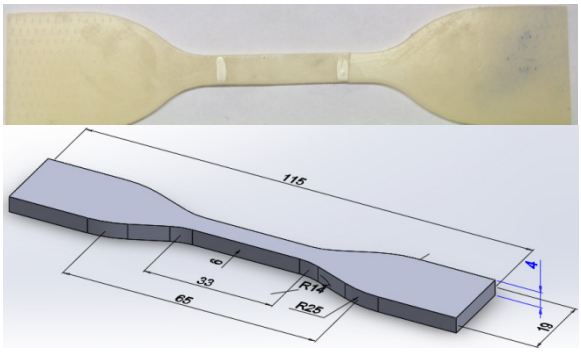


Figure 2: Tensile strength of test sample

The tensile properties were determined following the ASTM D638-02 procedure, using type IV test specimen dimensions (Figure 2) and using a Universal testing machine (Shimadzu Autograph AG-X Plus 20kN). This machine has the longitudinal stretch with high resolution camera (1.8 µm) and does not touch sample test. The crosshead speed was set at 50 mm/min at room temperature and four samples were tested for each composition.

The Izod impact strength was investigated according to ASTM D256. The composites were produced as 3.2x12.7x64 mm and a triangular with an angle of 45o, the radius of a glider at the bottom of the groove is r = 0.25 mm, shown in Figure 3. The experiment was performed with 5J collision energy and about 60 mm in length at room temperature. Five samples were tested for each composition.

Shore D scale was used to determine the hardness values of all samples. The tests were carried out the SHORE D Durometer DESIK. The durometer measures hardness by determining the depth of penetration into the material under test. The dial was graduated from 0-100 with a pointer sweep of 265º. Five samples of each formulation were tested and the average values were reported. The Fracture surface of each specimen in bending strength test was observed by Scanning Electron Microscope (HITACHI S - high resolution – 4800) with acceleration 5.0 kV. The surface of the samples used for SEM all was platinum-sputtered with a conductive layer before observation.

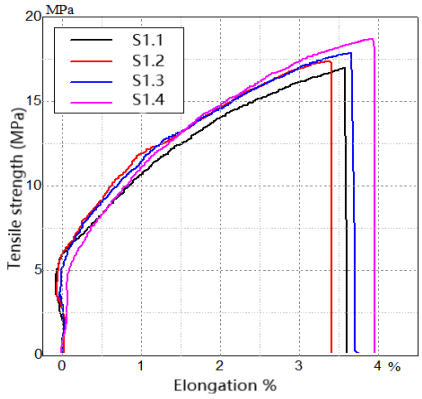


Figure 3: Sample measurement of impact resistance

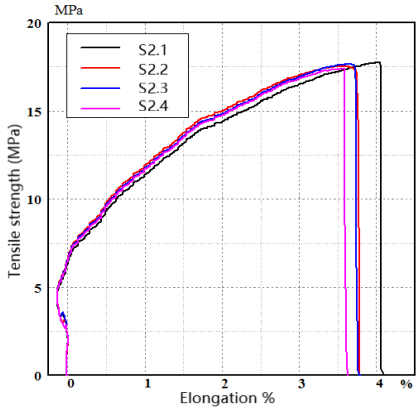
3. Result

3.1. Tensile strength

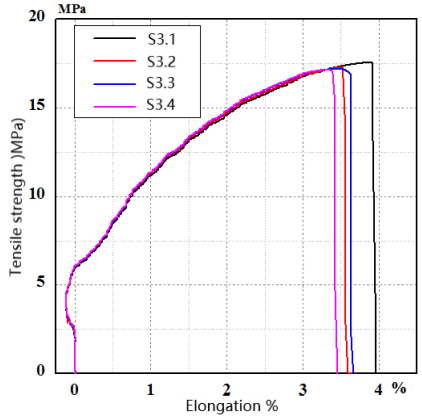
Figure 4 shows the stress-strain curve when pulling of the samples. Table 2 shows the tensile strength results of each sample group. From the data in Table 2, a chart showing the average value of the tensile strength of each sample group was established (Figure 5). The results of elongation at break in Figure 4 indicating the sample S4 at 3.50532% has the lowest elongation at break and sample S5 has the highest elongation at break (4.50412%). These values are an indication of the ductility of PP/LDPE/Kaolin blends. The higher value of elongation is showing the blend more ductile.



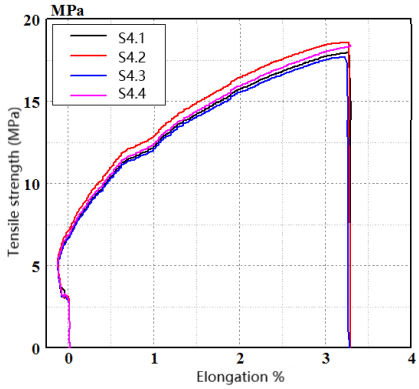
a) Sample S1 (0 wt.% kaolin content)



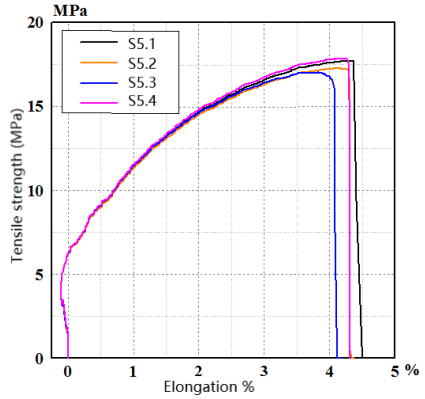
b) Sample S2 (3 wt.% kaolin content)



c) Sample S3 (5 wt.% kaolin content)



d) Sample S4 (10 wt.% kaolin content)



e) Sample S5 (14 wt.% kaolin content)

Figure 4: Stress-strain curve when pulling of the samples

Table 2: Results of tensile testing of samples

Sample	Average Tensile strength (MPa)	Average Strain (%)
S1	17.4904	3.69081
S2	17.5919	4.39359
S3	17.3139	3.68962
S4	18.1919	3.50532
S5	17.4757	4.50412

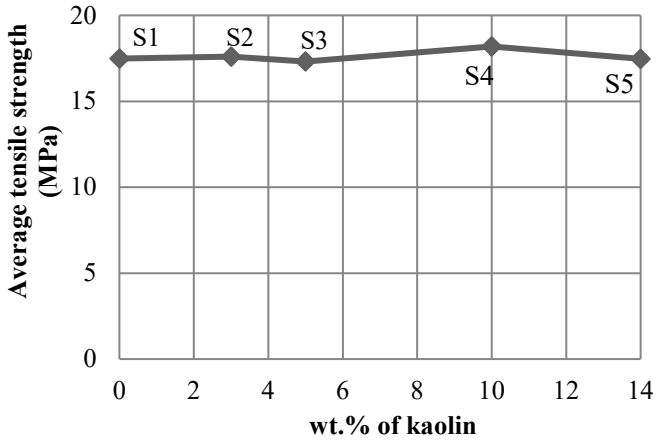


Figure 5: Average tensile strength of the samples

Figure 5 shows the average tensile results of the five samples. The tensile strength of PP/LDPE/Kaolin blends is increased with the addition of kaolin from 3% to 10 wt.%. 10% kaolin addition will have significant increase up to 18.1919 MPa comparing with the sample without kaolin (17.4904 MPa) while the higher amount of 14 % kaolin didn't seem to have better significant improvements. Tensile strength is an indication of the stiffness of a material. The adding of kaolin filler into the PP/LDPE matrix improves the stiffness of the blends. In the sample S4, the tensile strength is highest due to the distribution of filler on the PP/LDPE composite substrate surface that creates a hierarchical crystalline process between bridges. With sample S3, the tensile strength is lowest (17.3139 MPa) due to the large concentration of fillers in a given region which affects the crystallization. However, if there is more filler content in the polymer matrix will lead to the formation of the micro-filler and uneven kaolin particle size, and induces the difficulty of achieving a filling of filler in the PP/LDPE composite matrix forming holes [2]. This result influences the stress concentration at the boundary of the particle/ matrix and leads to a decreasing in particle/ matrix interactions. Therefore, it has a small effect on tensile strength [9].

3.2. Impact strength

The impact properties of PP/LDPE/Kaolin blends are summarized in Table 3 and Figure 6. As shown in the figure, the impact strength of PP/LDPE/Kaolin blends were decreased when increasing kaolin filler, 1.68806 kJ/m^2 for sample S1, 1.54280 kJ/m^2 for sample S2, 1.49422 kJ/m^2 for sample S3, 1.46986 kJ/m^2 for sample S4 and 1.30425 kJ/m^2 for sample S5. The reason is that PP/LDPE blends produce immiscible form because of the low interfacial adhesion. However, the immiscibility is good enough to preserve the good features of each polymer components of the blend. For example the impact strength of a polymer cannot be improved significantly by adding fillers with it [4]. In addition, in the mixture of PP/LDPE, kaolin acts as an evenly distribute effect in the PP/LDPE matrix and it will affect the chemical bonds between the kaolin and the PP/LDPE mixture [10]. For this reason, increasing kaolin filler content probably increased the level of stress concentration in the PP/LDPE composites with the resultant decrease in impact strength [9].

Table 3: Results of impact strength testing of samples

Sample	Average impact strength (kJ/m^2)
S1	1.68806
S2	1.54280
S3	1.49422
S4	1.46986
S5	1.30425

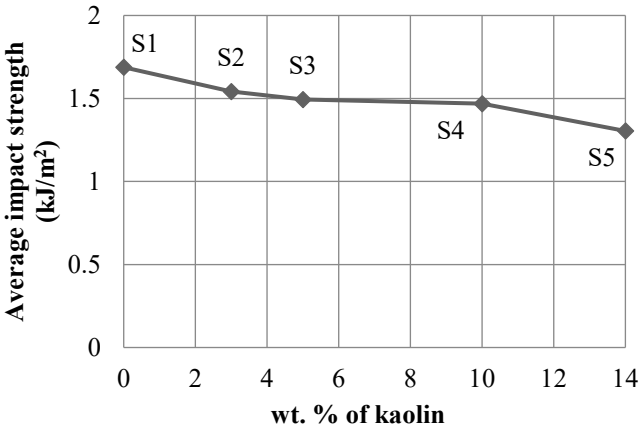


Figure 6: Impact strength of the samples

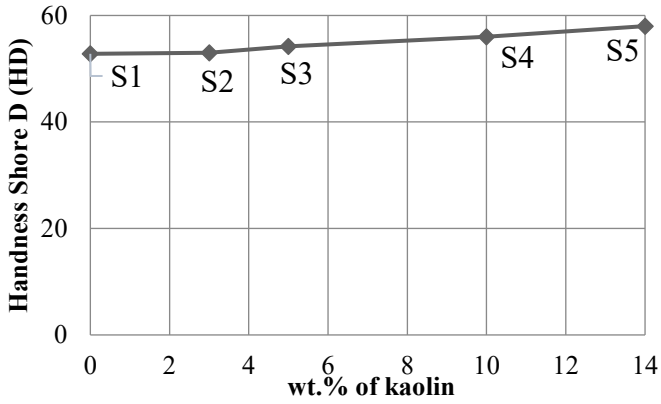


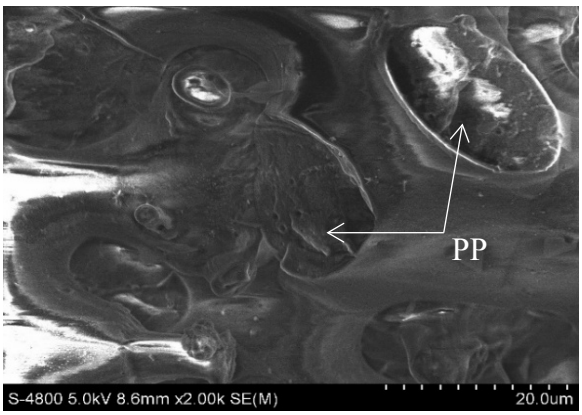
Figure 7: Hardness of the samples

3.3. Hardness

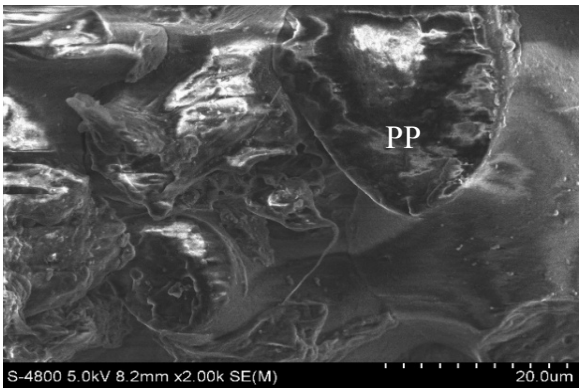
The hardness of PP/LDPE/Kaolin blends are showed in Figure 7. The hardness of PP/LDPE/Kaolin blends were all higher than the unfilled sample. The tests of hardness properties can evaluate the effects of adding kaolin fillers in PP/LDPE. It is evident from Figure 7 that a significant increase 52.8, 53, 54.2, 56 and 58 Shore D with adding 3, 5, 10, 14% kaolin, respectively. This is attributed to the fact that this filler acts as a reinforcing filler. Incorporation of the filler into the polymer matrix enhanced the stiffness of the material. The increase in hardness is due to the structure of the composite occurring in most reinforcement fillers. The higher the percentage of the filler incorporated, the harder the material, and the more rigid it becomes [10]. In composite PP/LDPE/kaolin, the mixture is characterized by the dispersion of vertical kaolin crystals in conjunction with the direction lines. When kaolin is combined with a thermoplastic mixture, it increases the hardness and produces concentrated stress, this also contributes to the reduction of the impact strength when added with kaolin filler [8].

3.4. Microstructure

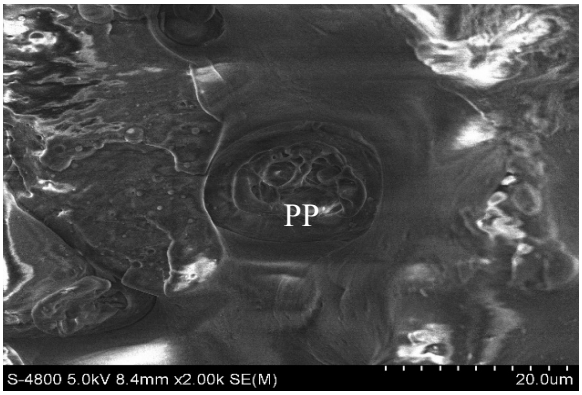
To better analyze the mechanical properties of the PP/LDPE/Kaolin mixture. SEM micrographs for the PP/LDPE/Kaolin blends were conducted. The results were shown in Figure 8. The sample S1 can see the pictures of the spherulites of PP in LDPE. PP is presumed to be the dispersed phase due to its high viscosity and elasticity. The PP spherulites size is coarse and could be easily distinguished. The reason is that PP and PE are compatible but only partially miscible. The PP/LDPE pairs tend to separate into two liquid phases. During the process of PP crystallization, the growth of PP spherulites in a "homogeneous" melt of a mixture of PE and PP will involve the propagating PP spherulite front encountering domains of PE melt [11]. When adding kaolin, PP spherulites is finer. The increasing the content of kaolin, the finer of the PP spherulites also increases.



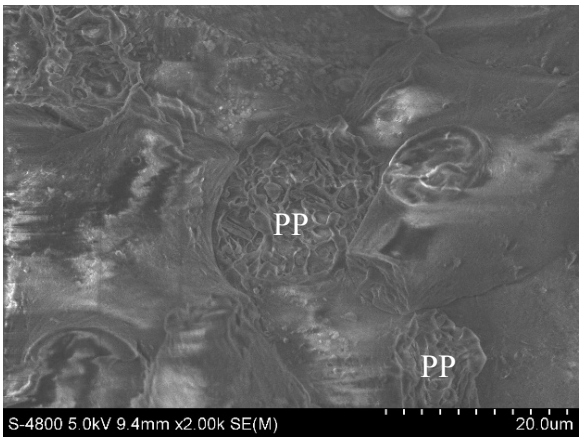
a) Sample S1 (0 wt.% kaolin content)



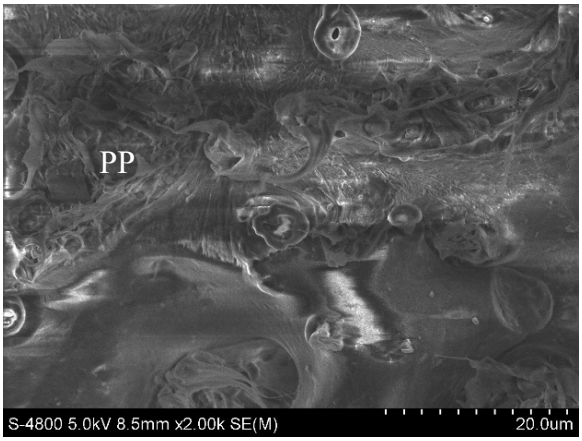
b) Sample S2 (3 wt.% kaolin content)



c) Sample S3 (5 wt.% kaolin content)



d) Sample S4 (10 wt.% kaolin content)



e) Sample S5 (14 wt.% kaolin content)

Figure 8: Surface collapse microstructure

4. Conclusion

From the above results, it was found that The tensile strength of PP/LDPE/Kaolin blends is increased with the addition of kaolin from 3% to 10 wt.%, but higher contents of kaolin 14% decreased the tensile strength.

When adding kaolin with PP/LDPE blend to increase the hardness but at the same time, it creates concentrated stress that reduces the impact strength. Adding kaolin to the PP/LDPE blend composition reduced the impact strength from 1.68806 kJ/m^2 to 1.30425 kJ/m^2 when increasing kaolin filler. On the contrary, the hardness of the PP/LDPE/Kaolin blends are increased by increasing the content of kaolin from 52.8 Shore D in samples without kaolin fillers to 58 Shore D in sample adding 14 wt.% kaolin filler.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

We are acknowledging HCMC University of Technology and Education, Dong Nhan Phat Co., Ltd and Material Testing Laboratory (HCMUTE). They gave me an opportunity to join in their team, accessed to the laboratory and research machines. Without their appreciated support it would not be possible to conduct this research.

References

[1] P. T. H. Nga; T. N. Thien, "Effect of Kaolin on the Mechanical Property of Polypropylene/Low Density Polyethylene Blend", 4th International Conference on Green Technology and Sustainable Development (GTSD), HCM City, Vietnam, 2018. DOI: 10.1109/GTSD.2018.8595576.

[2] M. Guessoum, S. Nekkaa, F. Fenouillot-Rimlinger, N. Haddaoui, "Effects of Kaolin Surface Treatments on the Thermomechanical Properties and on the egradation of Polypropylene", Hindawi Publishing Corporation International Journal of Polymer Science, 2012, 1-9, 2012. http://dx.doi.org/10.1155/2012/549154.

[3] H. Essabir, M. O. Bensalah, D. Rodrigue, D. Rodrigue, R. Bouhfid, A. e. k. Qaiss, "A comparison between bio- and mineral calcium carbonate on the properties of polypropylene composites", Construction and Building Materials, 134, 549-555, 2017. https://doi.org/10.1016/j.conbuildmat.2016.12.199.

[4] L. F. Kadhim, Z. F. Kadhim, "Studying The Properties of PP/LDPE polymer blend", Journal of Babylon University/ Engineering Sciences, 25(1), 193-201, 2017.

[5] M.S. Hassanien, A.I. Seedahmed, "Mechanical and rheological properties of polypropylene (PP)/ linear low density polyethylene (LLDPE) blend filled with talc and calcium carbonate compositions", Int J Eng Sci Res Technol, 4(1), 383-387, 2015.

[6] K. Şirin, M. Balcan, F. Doğan, The Influence of Filler Component on Mechanical Properties and Thermal Analysis of PP-LDPE and PP-LDPE/DAP Ternary Composites: Polypropylene, InTech China, 2012.

[7] M. Thomas, A. D. Kamble, N. John, "A Study on the Influence of Compatibilizer and Mica Filler on the Properties of Thermoplastic Polyurethane/Polyolefins Blends", Chem Sci Trans., 2(1), 181-191, 2013. DOI: 10.7598/cst2013.286.

[8] E.M. Khalaf, S.A. Awad, "Improvement of Mechanical and Water Absorbance Properties of Low-Density Polyethylene (LDPE) by using White Kaolin Powder (WK)", J. Adv. Chem. Sci., 3(1), 426-427, 2017.

[9] A. Mallika, A. K. Barik, B. Pal, "Comparative studies on physico-mechanical properties of compositematerials of low density polyethylene and raw/calcined kaolin", Journal of Asian Ceramic Societies, 3, 212-216, 2015. http://dx.doi.org/10.1016/j.jascer.2015.03.001.

[10] C. Onuoha, O.O. Onyemaobi, C.N. Anyakwo, G.C. Onuegbu, "Effect of filler loading and particle size on the mechanical properties of periwinkle shell filled recycled polypropylene composites", Am J Eng Res, 6(4), 72-79, 2017.

[11] R.A. Shanks, J. Li, L. Yu, "Polypropylene–polyethylene blend morphology controlled by time–temperature–miscibility", Polymer, 4(1), 2133-2139, 2000. https://doi.org/10.1016/S0032-3861(99)00399-7.

[12] S. N. Mustafa, "Effect of kaolin on the mechanical properties of polypropylene/ polyethylene composite material", Diyala Journal of Engineering Sciences, 5(2), 162-178, 2012.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 2, 527-536 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

A Simple Modelling Tool for Fast Combined Simulation of Interconnections, Inter-Symbol Interference and Equalization in High-Speed Serial Interfaces for Chip-to-Chip Communications

Davide Menin[*,1], Thomas Bernardi[1], Alessio Cortiula[1], Martino Dazzi[1,2], Alessio De Pr[1], Mattia Marcon[1], Marco Scapol[1], Andrea Bandiziol[3], Francesco Brandonisio[3], Andrea Cristofoli[3], Werner Grollitsch[3], Roberto Nonis[3], Pierpaolo Palestri[1]

[1]*University of Udine, Polytechnic Department of Engineering and Architecture, Via delle Scienze 206 33100 Udine, Italy*

[2]*Now with IBM Zurich, Switzerland*

[3]*Infineon Technologies Austria, 9500 Villach, Austria*

ARTICLE INFO

Article history:
Received: 14 January, 2020
Accepted: 20 February, 2020
Online: 10 April, 2020

Keywords:
Chip-to-chip communications
Modelling
Equalization

ABSTRACT

We describe an efficient system-level simulator that, starting from the architecture of a well-specified transmissive medium (a channel modelled as single-ended or coupled differential microstrips plus cables) and including the system-level characteristics of transmitter and receiver (voltage swing, impedance, etc.), computes the eye diagram and the bit-error rate that is obtained in high-speed serial interfaces. Various equalization techniques are included, such as feed-forward equalization at the transmitter, continuous-time linear equalization and decision-feedback equalization at the receiver. The impact of clock and data jitter on the overall system performance can easily be taken into account and fully-adaptive equalization can be simulated without increasing the computational burden or the model's complexity.

1 Introduction

This paper extends the work presented at the 42nd International Conference on Information and Communication Technology, Electronics and Microelectronics (MIPRO 2019) [1] and describes a simple and efficient tool for fast system-level simulation of high-speed serial interfaces, a topic that has received much attention in the past two decades due to its relevance in modern electronic systems: As the miniaturization of CMOS integrated circuits (IC) keeps following the path described by Moore's law [2], the amount of components integrated onto single devices, the number of functionalities available on single ICs and their speed increases significantly every few years. Over the last decades, evidence has arisen that the major bottleneck in performance shifted from computational capabilities and the associated power consumption towards communication between different ICs [3]. In fact, many applications such as modern microprocessors, servers, micro-controllers, FPGAs, even portable devices and, recently, automotive systems require High-Speed I/O (HSIO) modules capable of handling data rates up to 128 Gb/s with energies per bit as low as 1 pJ [4–9]. Moreover, as in many applications chip area and pin availability pose strict design constraints, the aforementioned devices cannot support parallel I/O that would reduce the data rate of individual channels, implying that such communications need to be implemented as high-speed serial interfaces (HSSI). In the present paper, the terms HSIO and HSSI will be used interchangeably to denote high-speed serial communication devices.

At multi-Gb/s data rates, performances are highly affected by impedance discontinuities in the interconnections due to PCB characteristics, presence of vias and package features; non-perfect impedance matching due to fabrication imperfections or poor compatibility for different devices; and by the dispersive nature of the transmissive medium at high frequencies [4, 10]. All these phenomena concur in causing Inter-Symbol Interference (ISI), which manifests itself as a smoothing and widening of the pulses sent along the channel so that they are superimposed to other symbols transmitted in the neighbouring unit intervals (UIs), thus increasing the Bit-Error Rate (BER) at the receiver, dramatically impairing the quality of the transmission [4, 10, 11].

In order to cope with this, HSIOs are required to implement complex equalization strategies, both at the transmitter or at the receiver,

[*]Davide Menin, Via delle Scienze 206, 33100 Udine (Italy), menin.davide@spes.uniud.it

and in the analog, mixed-signal or digital domains [10, 12–15]. Such techniques include Feed-Forward Equalization (FFE) at the transmitter, Continuous-Time Linear Equalization (CTLE) and Decision-Feedback Equalization at the receiver [10]. FFE uses an FIR filter that applies a pre-distortion to the transmitted pulses in order to preemptively compensate for the channel distortion; CTLE comprises a peaking amplifier mainly employed to compensate for the high frequency attenuation of the channel and possibly provide additional gain control at low frequency; in DFE the history of recent received bits is stored into a shift register and used to correct the received analog signal in order to cancel ISI at the input of the slicer, either through FIR or IIR filters.

One of the main challenges in the design and implementation of such equalization techniques is the fact that HSIOs are supposed to operate on a variety of channels whose features are unknown at design time. Thus, the optimal parameters of the equalizers cannot be precisely known and set a priori during the design phase, unless the resulting suboptimal performance can be tolerated, when it does not completely impede communication. Even in such rare cases where the transmissive medium is well known, the design itself is intrinsically dependent on process, voltage and temperature (PVT) variations and technology corners, all of which need to be counteracted by the equalizers. Therefore, calibration and adaptation strategies are required in order to find the optimal equalization parameters for the actual channel [13, 15–17]. Full adaptation automatically performs such a task, and is usually implemented in the form of Sign-Sign Least-Mean Squares (SS-LMS) algorithms due to the short time required to adjust the equalization parameters and the simplicity of their realization [11, 12, 15, 16, 18–21].

Moreover, HSIOs are also equipped with algorithms for clock and data recovery (CDR), and even performance monitoring, hence making up very complex electronic systems [12, 22]. Such a complexity cannot be conveniently handled through transistor-level descriptions because of the extremely long simulation times that they require. Therefore, various system-level models have been proposed in the last decades to aid the design of HSSIs, mainly using statistical techniques [19, 23–26]. Such tools are very important for the initial system-level assessment in the design of chip-to-chip HSIOs for selecting design specifications such as the number of equalization taps, the amount of high-frequency content that needs to be equalized, evaluating the Signal-to-Noise Ratio (SNR) and the overall jitter that can be tolerated without degrading the BER.

The design and analysis of HSIOs comprising a variety of complex equalization techniques require efficient system-level models capable of producing fast and accurate predictions of the system behaviour. Extending the work presented in [1] and relating it with contributions from [19, 27], this paper shows how fast system-level simulations of high-speed serial interfaces can be performed with a simple modular model. The paper proceeds as follows: Section 2, starting from the architecture of a generic HSSI, describes the numerical model, how it evaluates performance accounting for jitter and how fully-adaptive equalization is computed; Section 3 shows some sample simulation results and comparisons with post-layout transistor-level simulations, demonstrating the capabilities of the proposed approach; finally, conclusions are drawn in Section 4.

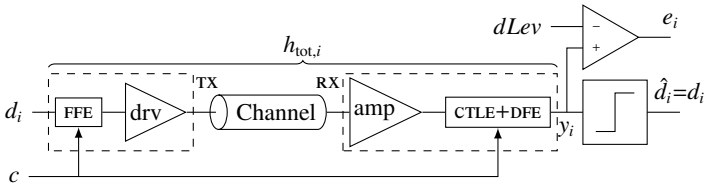


Figure 1: Scheme of a generic high-speed serial interface with equalization: $h_{\text{tot},i}$ is the overall pulse response of the TX (FFE+driver) + channel + RX (amplifier+CTLE+DFE) system; c is a generic equalization parameter (e.g. a filter tap), which may be either statically set or automatically adapted [19].

2 Model Description

2.1 Architecture of the Transceiver

In order to accurately model the system performance of a generic HSIO device, the general model depicted in Figure 1 and extensively described in [19] is considered: Denoting by the subscript i the sampling instant $t_{\text{B}} = iT_{\text{B}}$ (where T_{B} corresponds to a bit period, i.e. one Unit Interval UI), the data sequence d_i is sent by the differential transmitter TX at a bitrate $f_{\text{B}} = 1/T_{\text{B}}$, optionally implementing FFE; the channel, whose sampled pulse response is $h_{\text{ch},i}$, can be modelled either as two independent single-ended lines or as a coupled differential line; the receiver RX contains an amplifier, a CTLE and a DFE, and produces the analog voltage y_i; the slicer makes decisions on such a voltage ($\hat{d}_i = \text{sign}(y_i)$, i.e. $\hat{d}_i = 1$ if $y_i > 0\,\text{V}$ and -1 otherwise) and its sampling point can be modified with respect to the one determined by the CDR in order to perform optimal sampling [12, 19]; moreover, when performing full adaptation, the analog voltage y_i is compared with a reference voltage *dLev* [15] to determine the error e_i (i.e. the distance of the actual sample to *dLev*, usually defined as the desired voltage level corresponding to a '1' bit), and use this information to perform adaptation. Assuming that the BER is small (either because the channel has low loss or because it is well equalized), the reconstructed data $\hat{d}_i$ is equal to the transmitted data d_i.

2.2 Numerical Model of the Transceiver

The numerical model implemented in Matlab exploits a fast approach for the modelling of ISI and equalization in HSIOs, the flowchart of which is here summarised in Figure 2 and detailed in the following paragraphs.

2.2.1 Transmitter

The idealised transmitted waveform is modelled as the trapezoidal pulse $v_{\text{pulse}}(t)$, shown in Figure 3a and characterised by its duration, amplitude and by the slope of its edges (rise and fall times $t_{\text{rise}} = t_{\text{fall}}$); such a waveform is easily Fourier-transformed, giving $V_{\text{pulse}}(f)$, to which FFE is applied by summing weighted delayed versions of the transformed pulse itself:

$$V_{\text{TX}}(f) = V_{\text{pulse}}(f) \sum_{n=0}^{N_{\text{FFE}}-1} w_n e^{-i\omega n T_{\text{B}}}, \tag{1}$$

where w_n are the weights of the N_{FFE} FFE taps, subject to the constraint $\sum_{n=0}^{N_{\text{FFE}}-1} |w_n| = 1$ due to the fact that the power available in the

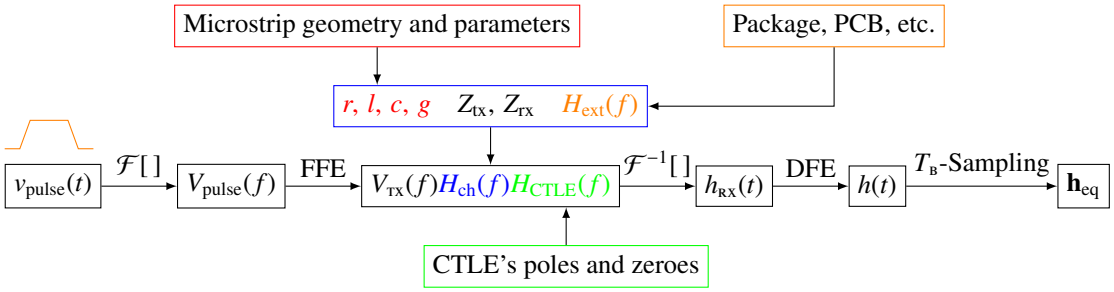


Figure 2: Diagram of the procedure implemented to obtain the received pulse response [1]. $\mathcal{F}$ and $\mathcal{F}^{-1}$ stand for the Fourier and the inverse Fourier transforms, respectively.

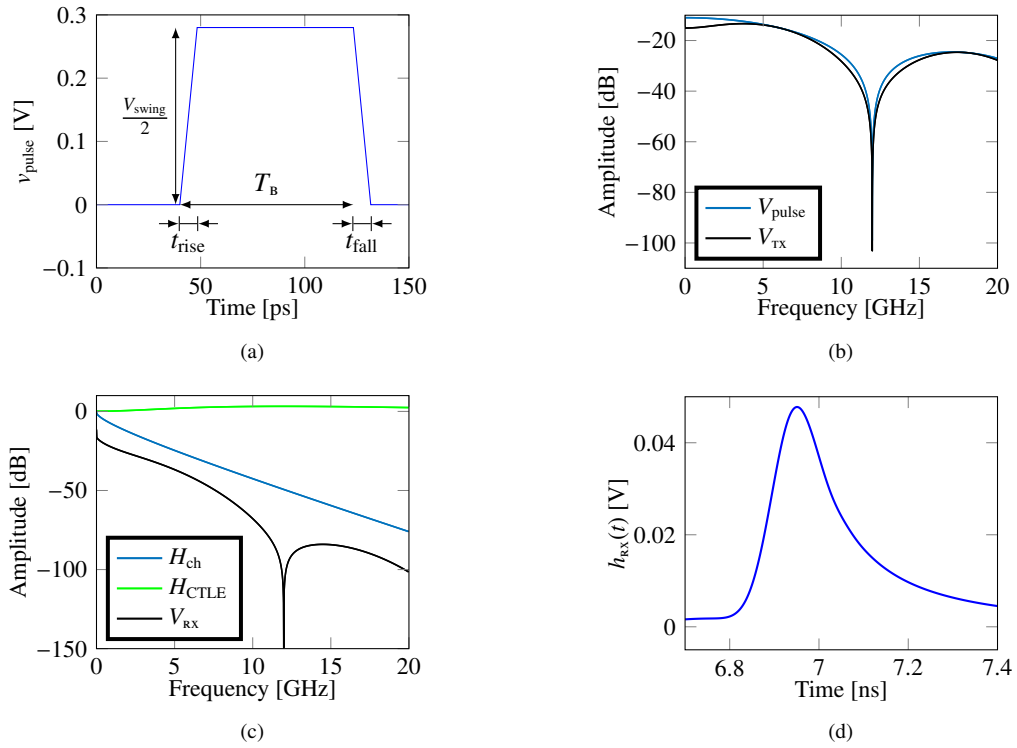


Figure 3: Example based on a PCIe 4.0 channel attenuating 20 dB at 6 GHz, showing the waveforms elaborated by the fast modelling tool according to the flow of Figure 2 [1, 19]. (a) shows the transmitted trapezoidal pulse $v_{\text{pulse}}(t)$; (b) reports the corresponding Fourier transform prior to ($V_{\text{pulse}}(f)$) and after application of FFE ($V_{\text{TX}}(f)$); (c) shows $H_{\text{ch}}(f)$, $H_{\text{CTLE}}(f)$, and the result of their multiplication with $V_{\text{TX}}(f)$; (d) shows the subsequent inverse Fourier transform prior to DFE $h_{\text{RX}}(t)$, the application principle of which is shown in Figure 5.

driver is limited. The effect of such an operation is shown in the frequency domain in Figure 3b.

The proposed approach considers a transmitter's impedance which is kept constant and does not change at high or low outputs, as is the case in Input-Output Buffer Information Specification (IBIS) models [28, 29]. Moreover, the pulse shape used in the proposed method (trapezoidal shape with $t_{\text{rise}} = t_{\text{fall}}$) is chosen in order to exploit the channel's linearity and hence use the channel's pulse response in the various computations instead of its step response [26]; in other words, a sequence of pulses $v_{\text{pulse}}(t)$ having the same amplitude results in a constant voltage level, which is not the case with other pulse shapes, see [30].

2.2.2 *Communication Channel*

The transmissive medium is generally modelled as a differential line, made up either of two independent microstrips (to simulate lines placed at some distance from each other as to minimise interactions) or a single coupled microstrip excited with an odd mode (to realistically reproduce differential signalling). This is then used to reproduce the salient features of any other type of transmission line, such as the target attenuation at a certain frequency or its characteristics impedance.

The microstrip features are computed from the line's geometry and material parameters following the approach defined in [31–36] for single-ended lossy microstrips, or extended to the coupled case according to [37, 38], and then used to extract its per-unit-length parameters $r(f)$, $l(f)$, $c(f)$, $g(f)$ considering dielectric losses and skin effect, all of which are among the main contributors to ISI [4].

Such a result can then be combined with $H_{\text{ext}}(f)$, a transfer function representing the socket or package and incorporating notch filters, which can be used to take into account discontinuities, vias, etc. in order to provide a complete description of realistic channels.

$H_{ext}(f)$ can be calculated with a model of the package, e.g. in terms of parasitic resistance, inductance and capacitance, which allows a straightforward evaluation of its transfer function in terms of poles and zeroes, while the contributions due to impedance discontinuities or vias can be taken into account by fitting the features of actual measurements of the transmission line's S parameters to the transfer function of notch filters in the form

$$H_{\text{notch}}(f) = \frac{1 + 2\xi(if)/f_0 + (if)^2/f_0^2}{1 + 2(1-\xi)(if)/f_0 + (if)^2/f_0^2}, \tag{2}$$

where ξ is the filter's damping factor and f_0 is its notch frequency.

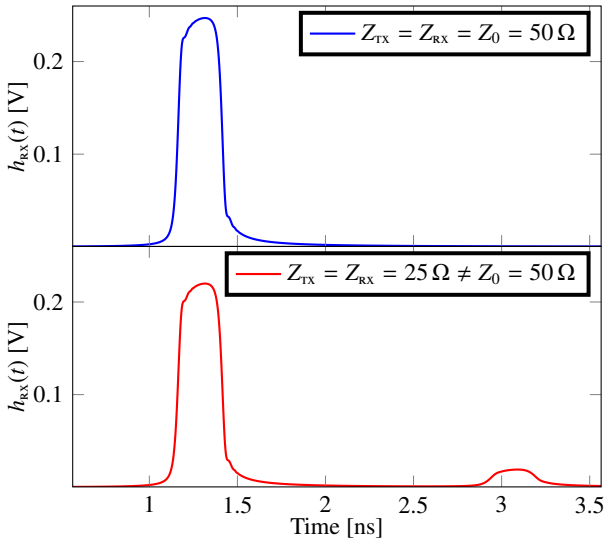


Figure 4: Effect of mismatch on the received pulse response: On top, $Z_{TX} = Z_{RX} = Z_0 = 50\,\Omega$; below, $Z_{TX} = Z_{RX} = 25\,\Omega$, $Z_0 = 50\,\Omega$, showing reflections due to such a mismatch. Transmission at 4 Gb/s over a low-loss channel (approximately −3 dB at Nyquist frequency).

Using as additional parameters the driver's and the receiver's termination impedances (Z_{TX} and Z_{RX}, respectively), the transmission line transfer function is computed from the telegrapher's equations as

$$H_{ch}(f) = \frac{e^{-\gamma L}(1+\Gamma_{RX})(1-\Gamma_{TX})}{1-\Gamma_{TX}\Gamma_{RX}e^{-2\gamma L}} H_{ext}(f), \tag{3}$$

where $\gamma = \sqrt{(r+i\omega l)(g+i\omega c)}$ is the propagation coefficient, L is the line length and $\Gamma_{TX/RX} = \frac{Z_{TX/RX} - Z_0(f)}{Z_{TX/RX} + Z_0(f)}$ are the reflection coefficients corresponding to the transmitter and the receiver. Note that, due to the inclusion of $Z_{TX/RX}$, (3) takes into account possible non-perfect matching among driver, transmission line and receiver, which is shown as an example in Figure 4: The mismatch produces a reflection that contributes to ISI.

2.2.3 Receiver

The CTLE is modelled as a rational function $H_{CTLE}(f)$ characterised by the CTLE's poles and zeroes; optionally, an extraction of the CTLE's transfer function from simulations of the transistor-level HSIO can be used to reproduce more accurately a realistic implementation (and the frequency response of other analog blocks in the receiver can be similarly taken into account). The received signal associated to the transmitted trapezoidal pulse has spectrum $V_{RX}(f) \triangleq V_{TX}(f)H_{ch}(f)H_{CTLE}(f)$; an example of this is shown in Figure 3c with some of its sub-components.

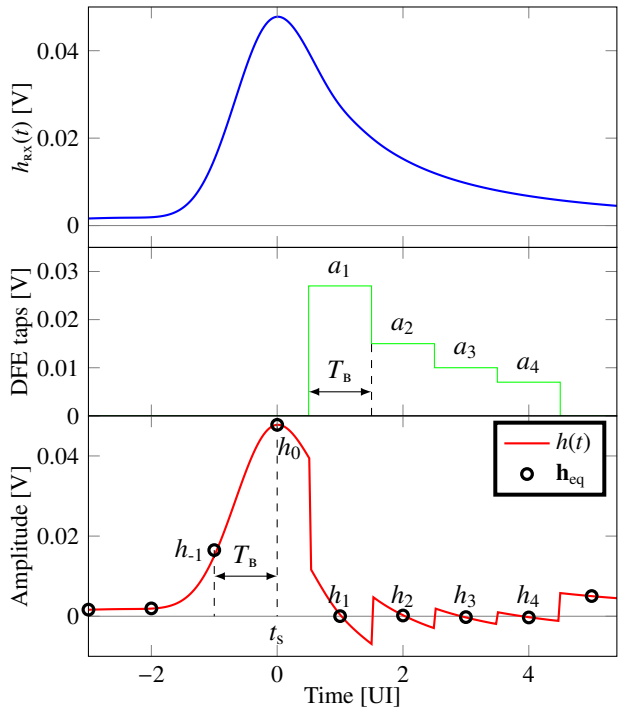


Figure 5: Procedure used by the model to apply the DFE correction to the received analog pulse response $h_{RX}(t)$ to obtain $h(t)$ and eventually its sampled version $\mathbf{h}_{eq}$. The DFE taps are rectangular pulses 1 UI wide and centred on the sampling point determined by the CDR.

The received pulse response $h(t)$ at the slicer's input is then obtained through inverse Fourier transform of $V_{RX}(f)$ using the procedure in [39], yielding $h_{RX}(t)$ (an example of $h_{RX}(t)$ is shown in Figure 3d). Application of the DFE correction is performed as shown in Figure 5, i.e. by subtracting from $h_{RX}(t)$ rectangular pulses with amplitude equal to the tap weights a_i and centred on the sampling point determined by the CDR.

The procedure above implicitly assumes that the CDR has reached its steady state and is locked. Its impact on the behaviour of the HSSI is twofold: It determines the sampling point for data, error and edge samples (which is related to the position of the "rectangles" associated to the DFE, as mentioned above), while the jitter at its output is responsible for a reduction of BER (as will be explained in Section 2.4). For what the sampling point is concerned, we can simply assume that the data samples correspond to the maximum of the pulse response $h(t)$; alternatively, an Alexander CDR [40] can be emulated by determining the time instants corresponding to $h_{RX,-0.5}$ and $h_{RX,0.5}$ (which are the positions of the edge samples of the CDR in a real implementation [27]) and then assume that the data sample is exactly in between. On top of it, an algorithm for optimal sampling point may be used to determine a shift from the output of the CDR, which results in an improved sampling position [12, 19]. Any of the above can be selected and all of them aim at sampling as close to the centre of the eye as possible in order to reduce the probability of error.

Sampling by the bit period T_B is eventually performed in order to

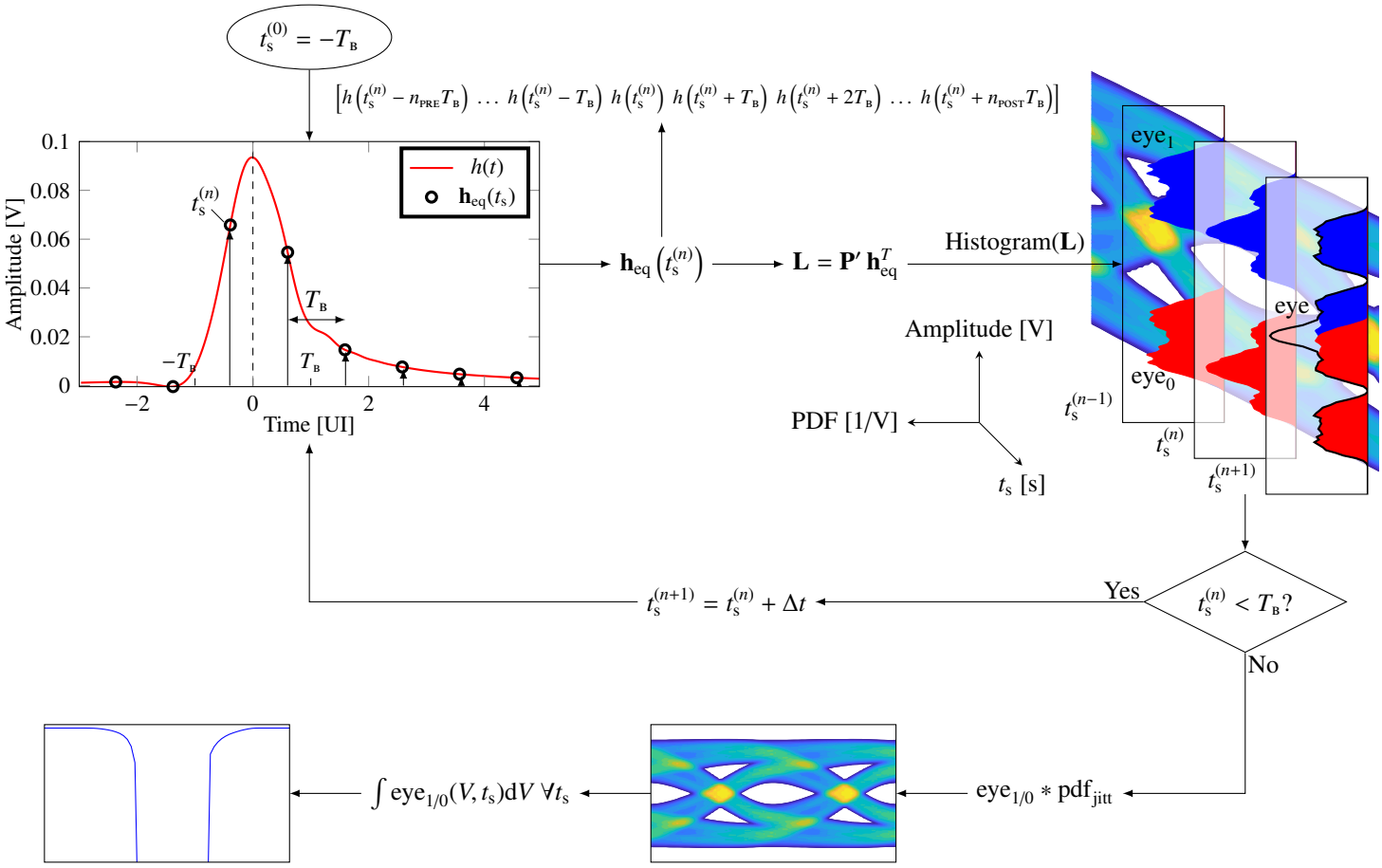


Figure 6: Diagram of the procedure used to evaluate the HSSI performance in terms of eye diagram and bathtub plot, as explained in Section 2.3. All plots and diagrams are obtained by transmitting at 12 Gb/s over a channel losing 14.3 dB at Nyquist frequency, with the FFE applying a de-emphasis of −2.5 dB; no DFE is applied for the sake of demonstrating the working principle of the proposed method. From an initial sampling point $t_s^{(0)} = -T_B$, at each iteration $n_{PRE} + n_{POST} + 1$ T_B-spaced samples of $h(t)$ are taken to form $\mathbf{h}_{eq}\left(t_s^{(n)}\right)$; at each sampling point, histograms of $\mathbf{L}$ are computed to determine the corresponding PDFs, represented in the drawing at the top right as cross sections of the eye diagram; the procedure is repeated until t_s reaches T_B and then the eye diagram and bathtub plot are computed taking into account jitter at the receiver.

obtain the analog voltages (vector $\mathbf{h}_{eq} = [h_{-n_{pre}} \ldots h_{-1}\, h_0\, h_1 \ldots h_{n_{post}}]$) that would be sensed at the input of the slicer at sampling time.

2.3 Evaluating the HSSI Performance

Performances of the transceiver are determined mainly by computing the eye diagram, constructed by folding the received signal over a time length of 1 UI, which allows to observe all the transitions that take place during operation of the serial link and their density; and by calculating the bathtub plot, which shows the cumulative distribution function of the received errors over the same time span of the eye diagram, indicating the sampling positions that result in an increased BER [4]. Both such metrics require probabilistic calculations in order to maintain computation times low [19, 23].

From the sampled pulse response $\mathbf{h}_{eq}$ one can compute all possible values of the voltage y_i at the samplers, due to all the possible sequences of bits that can be sent, as

$$\mathbf{L} = \mathbf{P}\,\mathbf{h}_{eq}^T, \tag{4}$$

where $\mathbf{L}$ is a column vector containing such voltage levels and $\mathbf{P}$ is a *permutation matrix* which contains all the possible bit sequences of a certain length that can be transmitted. In fact, $\mathbf{P}$ is structured as a truth table: It features a number of columns equal to $n_{PRE} + 1 + n_{POST}$, where n_{PRE} and n_{POST} are the number of pre- and post-cursors, respectively, which can be chosen according to their relevance in the pulse response; while the number of rows is $2^{n_{PRE}+1+n_{POST}}$, i.e. the number of all possible sequences composed of $n_{PRE} + 1 + n_{POST}$ bits. In other words, $\mathbf{L}$ considers all the possible ways in which the samples of the pulse response can combine due to ISI, hence simulating observation of the received analog voltage y_i over a sufficiently long time span. Moreover, $\mathbf{L}$ implicitly depends on the choice of the sampling instant t_s through the sampled pulse response $\mathbf{h}_{eq}$.

$$\mathbf{P}' = \begin{pmatrix} -1 & 1 & -1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}. \tag{5}$$

By assuming that the eye is vertically symmetric (the transceiver behaviour when transmitting a '1' bit is the same as though a '0' was sent, just with a sign reversal), only the cases in which $\hat{d} = 1$

are useful for the purpose of computing the HSIO performance. By coding the '1' and '0' bits as 1 and −1 values, respectively, and keeping only the non-redundant rows, e.g. for one pre- and two post-cursors such a reduced matrix (denoted by ') reads

Equation (4) provides an easy way to compute the eye diagram and the bathtub plot. The eye diagram can be computed by sampling $h(t)$ at various $t = t_s \in [-T_B/2, T_B/2]$, where T_B is the bit period (1 UI), and creating histograms $\text{eye}_1(V, t_s)$ of the corresponding $\mathbf{L}(t_s)$; due to the fact that the eye for the '0' bit is just the flipped version of the one for the '1' bit (as follows from the assumption of symmetry), they can be combined to obtain the overall eye diagram

$$\text{eye}(V, t_s) = \frac{\text{eye}_1(V, t_s) + \text{eye}_1(-V, t_s)}{2}. \tag{6}$$

The bathtub plot, i.e. the BER corresponding to a voltage threshold equal to 0 V as a function of the sampling instant t_s, is then given by the probability that a '1' bit is misinterpreted for a '0' (that is the same as the probability that a '0' bit is misinterpreted for a '1', due to the above assumption of symmetry):

$$\text{BER}(t_s) = \int_{-\infty}^{0} \text{eye}_1(V, t_s) \text{d}V. \tag{7}$$

The overall procedure for computing the eye diagram and the bathtub plot is shown in Figure 6, summarising the flow described in this Section and in the following.

2.4 Including the Effect of Jitter

The effect of jitter on the receiver can be optionally taken into account by simply convolving the single $\text{eye}_{1/0}$ of (6) and the probability density function of sampling time t_s corresponding to the jitter component of interest $\text{pdf}_x(t_s)$ [41]. As an example, considering an oscillator in the receiver affected by random jitter, the period jitter of which has variance σ_{PJ} (in other words, a clock with phase noise going as $1/f^2$, which means that the jitter values in different periods are uncorrelated) and a CDR having a bandwidth BW_{CDR}, the squared variance of the absolute jitter present in the recovered clock can be easily shown to be given by

$$\sigma_{RJ}^2 = \frac{\sigma_{PJ}^2}{4\pi T_B BW_{CDR}}. \tag{8}$$

The simple example shown in Figure 7 assumes a jitter characterised by a Gaussian distribution described as

$$\text{pdf}_G(t_s) = \frac{1}{\sqrt{2\pi}\sigma_{RJ}} e^{-\frac{1}{2}\left(\frac{t_s}{\sigma_{RJ}}\right)^2}, \tag{9}$$

where σ_{RJ} is the variance of the random jitter affecting the recovered clock as per (8), which in general may include other sources than random jitter of the clock alone.

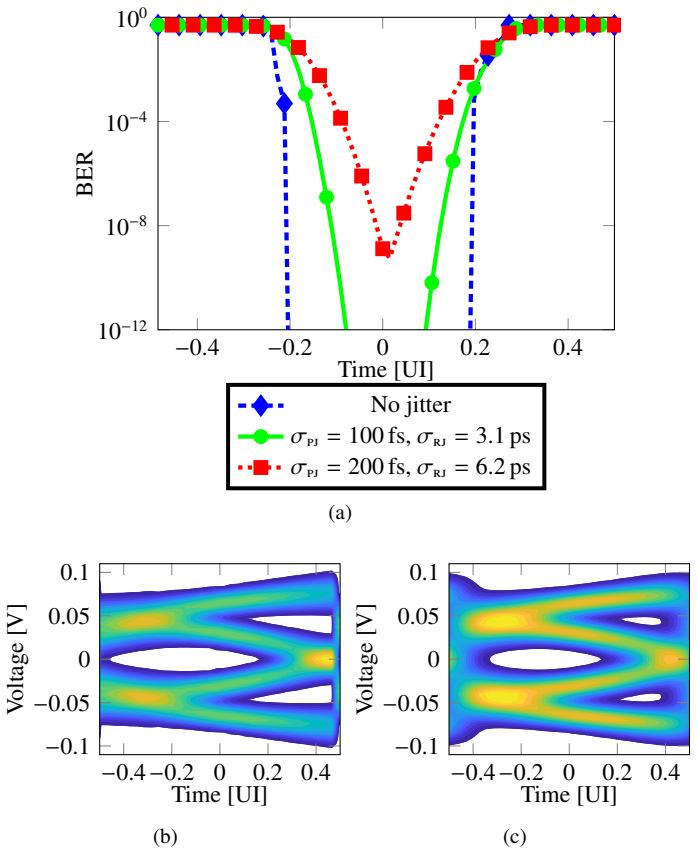


Figure 7: PCIe 4.0 channel attenuating 20 dB at 6 GHz; approximately 4 dB of FFE de-emphasis, first 4 post-cursors cancelled by the DFE. The random jitter applied at the receiver's recovered clock is calculated according to (8) assuming a CDR bandwidth BW_{CDR} = 1 MHz. (a) BER bathtub plots; eye diagrams with (b) no random jitter and (c) σ_{RJ} = 6.2 ps random jitter on the recovered clock.

2.5 Including Fully-Adaptive Equalization

As briefly mentioned in the Introduction, the problem concerning the optimal settings of the equalizers' parameters is not trivial, and often one must resort to full adaptation in order to automatically find optimal equalization parameters. The implementation of fully-adaptive techniques based on an SS-LMS algorithm in the simulation approach described in this paper is relatively straightforward and was thoroughly described in [19]: Briefly, it involves computing quantities in the form

$$c^{(k+1)} = c^{(k)} + \mu_c \text{sign}\left(\hat{d}_i\right) \text{sign}\left(e_j\right), \tag{10}$$

where μ_c is the step size, $\hat{d}$ is the data sample received at time i, e_j is the error at time $j = i + k$ between the analog voltage y_j at the samplers and the desired voltage $dLev$ corresponding to a '1' bit and $c^{(k)}$ is the k-th iteration on a generic parameter that can be adapted (the taps amplitude of FFE or DFE, the positions of poles and zeroes in a CTLE modelled e.g. as $H_{CTLE}(s) = c_0 + c_1 s$, the sampling phase as well as $dLev$ itself). Note that $\text{sign}\left(\hat{d}_i\right) \text{sign}\left(e_j\right)$ corresponds to correlating the error made at a certain time with the bit received at possibly another time in the past or in the future, where obviously the latter can be considered only when data and errors are parallelized before computation of the fully-adaptive algorithm [8, 17, 27, 42–44]. Such correlations provide information

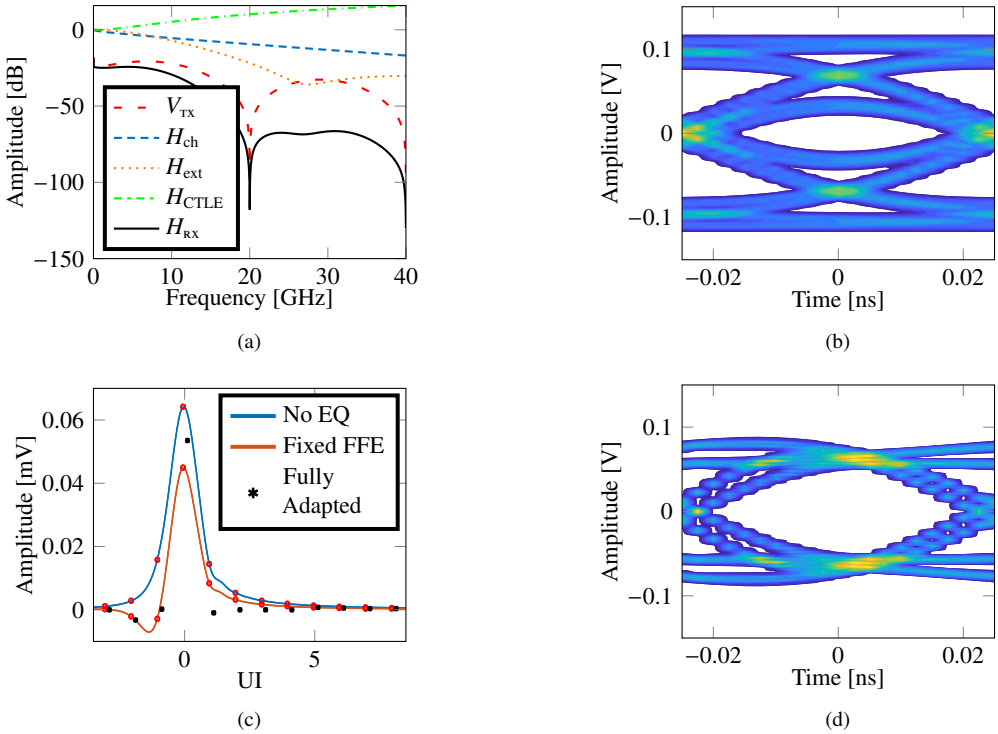


Figure 8: Simulation of a channel with 12 dB attenuation at 10 GHz and fixed 2-tap FFE. All the relevant Fourier transforms of signals and transfer functions of the transceiver are plotted in (a); the eye diagram prior to equalization is shown in (b); (c) shows the pulse response prior to equalization (solid blue line), after the fixed FFE (solid orange line) and after application of the fully-adaptive loop (Figure 10) and subsequent sampling (black asterisks); (d) shows the resulting equalised eye diagram.

on whether to increase or decrease the corresponding parameter c and bring it to convergence, and are usually collected and averaged over time [12, 27]. Such a correlation can be easily evaluated from $\mathbf{h}_{eq}$ by multiplication with a matrix similar to $\mathbf{P}'$ of (5) (further mathematical details are given in [19]).

By applying equalization to the pulse response of the transceiver plus channel, as mentioned in Section 2.2, iteration of (4) and Equation 10 provides the evolution over time of the adaptation procedure until the equalization parameters converge in the neighbourhood of the optimum [19].

3 Results

As an example of the power and versatility of the proposed approach, we consider here the fully-adaptive equalization of an HSIO transmitting at 20 Gb/s with rise/fall times equal to 20 % of the UI and ±0.25 V differential voltage swing on a channel attenuating 12 dB at 10 GHz. A fixed 2-tap FFE with a pre-emphasis of approximately 6 dB ($w_{-1} = -0.25$, $w_0 = 0.75$) was applied at transmit side to reduce the first pre-cursor. The package was described by an LC π-network with $L = 2$ nH and $C = 100$ fF, while an impedance discontinuity was modelled by adding a notch filter with $\xi = 0.1$ centred at $f_0 = 27$ GHz; both features were included through the function $H_{ext}(f)$ mentioned in Section 2.2.2. Figure 10 shows the resulting simultaneous adaptation of *dLev*, CTLE's zeroes, DFE taps and sampling phase as a function of the number of iterations performed by the fully-adaptive algorithm described above. As expected, *dLev* converges to the average value corresponding to the '1' bit (i.e. the peak value of $h(t)$: h_0) and, as it approaches such a value, the other equalization parameters start to adapt and eventually converge: The DFE taps reach the values of the corresponding post-cursors (h_1, h_2 and h_3), while the phase is shifted w.r.t. the position determined by the CDR to a value that zeroes the first pre-cursor. Frequency representations of all relevant signals and transfer functions of the transceiver are plotted in Figure 8a, the unequalised eye diagram is depicted in Figure 8b, the channel pulse response is shown in Figure 8c prior to equalization, after the fixed FFE and after full adaptation, a situation to which corresponds the eye diagram of Figure 8d.

In order to validate the proposed numerical model, a comparison was carried out in [27] between the eye diagram obtained with the model itself and that obtained through post-layout transistor-level simulations. An HSSI for automotive applications implemented in 28 nm planar CMOS technology was simulated at 12 Gb/s at transistor level with full adaptation enabled; the numerical model was then used as a comparison in terms of performance and behaviour of the SS-LMS adaptive algorithm when the HSIO was used to communicate over a realistic, high-loss channel (−33 dB at 6 GHz) representing a transmission line as will likely be defined by the MIPI A-PHY standard. The results of Figure 9 show a good degree of accuracy in reproducing the transistor-level simulations when relevant features of the post-layout transistor-level implementation (chiefly, the transfer functions of the CTLE and of the Variable-Gain Amplifiers in the receiver) were extracted and used in the tool, as explained in Section 2.2.3.

As a means of comparison, in order to observe convergence of

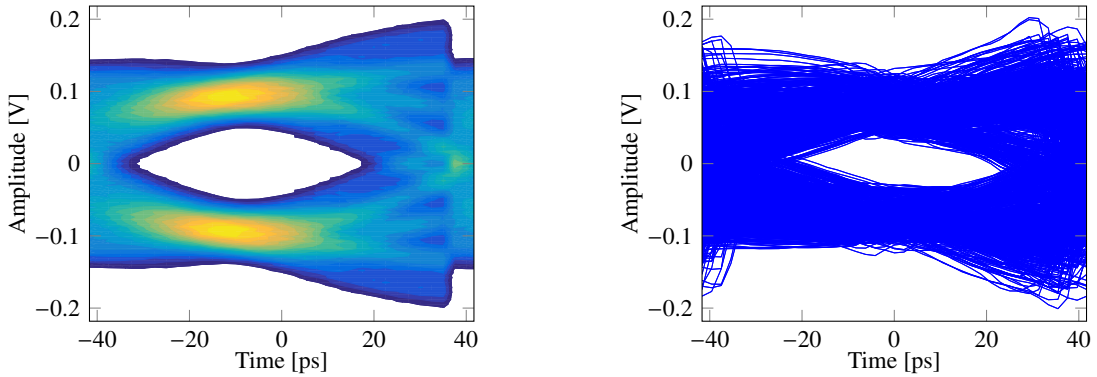


Figure 9: Eye diagrams after convergence of the SS-LMS equalization loops on a MIPI A-PHY channel [27] as obtained through (a) post-processing of the pulse response with the proposed method and (b) transistor-level simulations.

the fully-adaptive algorithm and be able to compute an eye diagram containing enough UIs, transistor-level simulations run at a speed of about 6 h per µs of simulation for at least 1 µs, whereas the proposed method takes about 5 s to provide the results.

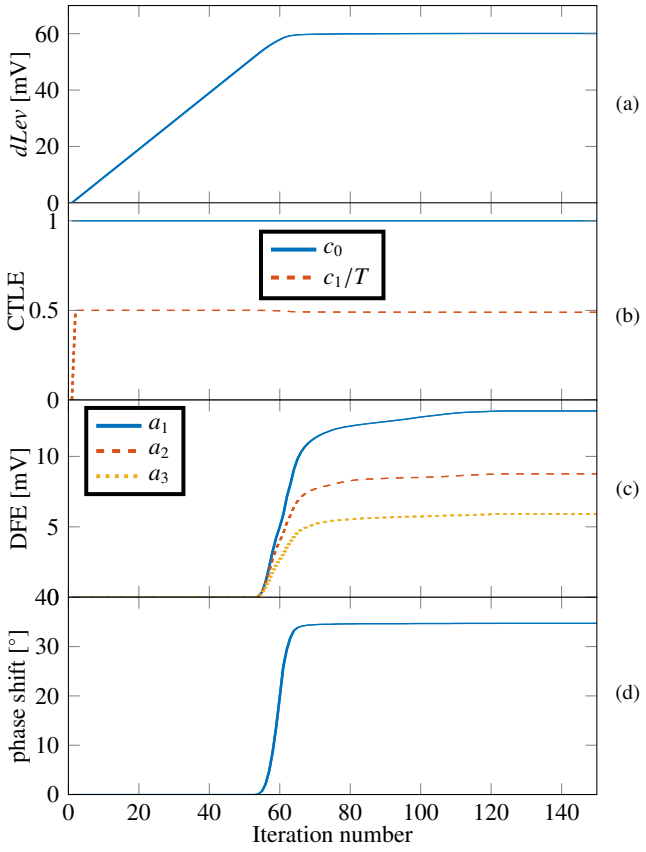


Figure 10: Simulation of a channel with 12 dB attenuation at 10 GHz and fixed 2-tap FFE; the trapezoidal pulse was sent at 20 Gb/s bitrate, it has 20 % rise/fall time and ±0.25 V swing differential voltage. From top to bottom, (a) *dLev* drives adaptation of (b) first-order CTLE , (c) 3-tap DFE and (d) optimal sampling point.

Preliminary tests employing behavioural models for the HSSI indicate a simulation speed of about 7.3 s per µs of simulation (not shown).The above is meant to be just a rough comparison, mainly because the various models do not necessarily implement all the components of an HSIO (e.g. the transistor-level simulation does not consider the digital part of the system), but it still provides some figures to consider when dealing with such kind of simulations.

4 Conclusions

We have presented a fast tool exploiting a simple modelling approach to evaluate the performance of high-speed serial interfaces for chip-to-chip communications. An efficient probabilistic algorithm was developed to evaluate the eye diagram with low computational effort. Sharing the same motivations of other similar models developed in the past in the literature, such an approach represents a powerful alternative to time-domain simulations, since complex systems working with BER as low as 10^{-15} require simulating very large amounts of bit periods, which may be very time consuming. The effect of the most common equalization strategies, of CDR techniques, of jitter and of full adaptation of the equalizers can be easily included in the model, so that the proposed simulation approach can be used for the system-level assessment of high-speed interfaces that need to comply with various standards. As examples of the capabilities of the proposed approach, we reported results from two interfaces: One transmitting at 20 Gb/s over a relatively low-loss channel (−12 dB at Nyquist frequency) and another operating at 12 Gb/s over a high-loss MIPI A-PHY line (−33 dB at Nyquist frequency). Both cases show that the combination of various equalization techniques is required to obtain suitable BERs, and that the proposed approach provides results that are comparable with much longer, time-domain post-layout transistor-level simulations, thus demonstrating the power of our model to evaluate the performance of realistic high-speed serial interfaces.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgement The authors would like to thank prof. Luca Selmi (University of Modena and Reggio Emilia) for support.

References

[1] A. Cortiula, M. Dazzi, M. Marcon, D. Menin, A. Bandiziol, A. Cristofoli, W. Grollitsch, R. Nonis, and P. Palestri, "A simple and fast tool for the modelling of inter-symbol interference and equalization in high-speed chip-to-chip interfaces," in *The 42nd Int. Conv. on Information and Communication Technol., Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, May 2019, pp. 116–120. doi: 10.23919/MIPRO.2019.8756752.

[2] G. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. 38, no. 8, pp. 114–117, Apr. 1965.

[3] H. Tamura, "Looking to the Future: Projected Requirements for Wireline Communications Technology," *IEEE Solid-State Circuits Mag.*, vol. 7, no. 4, pp. 53–62, 2015. doi: 10.1109/MSSC.2015.2477017

[4] T. C. Carusone, "Introduction to Digital I/O: Constraining I/O Power Consumption in High-Performance Systems," *IEEE Solid-State Circuits Mag.*, vol. 7, no. 4, pp. 14–22, Fall 2015. doi: 10.1109/MSSC.2015.2476016

[5] Y. Chen, P.-I. Mak, L. Zhang, and Y. Wang, "A 0.002-mm^2 6.4-mW 10-Gb/s Full-Rate Direct DFE Receiver With 59.6Channel Loss at Nyquist Frequency", *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 12, pp. 31073117, Dec. 2014. doi: 10.1109/TMTT.2014.2360697.

[6] N. J. Endo, "Wireless Communication In and Around the Car: Status and Outlook. ES3: High-Speed Communications on 4 Wheels: What's in Your next Car?" in *2013 IEEE Int. Solid-State Circuits Conf. Dig. of Tech. Papers*, Feb. 2013, pp. 515–515. doi: 10.1109/ISSCC.2013.6487598.

[7] J. Kim et al., "A 112 Gb/s PAM-4 56 Gb/s NRZ Reconfigurable Transmitter With Three-Tap FFE in 10-nm FinFET," *IEEE J. Solid-State Circuits*, vol. 54, no. 1, pp. 29–42, 2019. doi: 10.1109/JSSC.2018.2874040

[8] J. Lee, K. Park, K. Lee, and D.-K. Jeong, "A 2.44-pJ/b 1.62–10-Gb/s Receiver for Next Generation Video Interface Equalizing 23-dB Loss With Adaptive 2-Tap Data DFE and 1-Tap Edge DFE," *IEEE Trans. Circuits Syst. II*, vol. 65, no. 10, pp. 1295–1299, Oct. 2018. doi: 10.1109/TCSII.2018.2846677

[9] Z. Toprak-Deniz, J. E. Proesel, J. F. Bulzacchelli, H. A. Ainspan, T. O. Dickson, M. P. Beakes, and M. Meghelli, "A 128-Gb/s 1.3-pJ/b PAM-4 Transmitter With Reconfigurable 3-Tap FFE in 14-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 1, pp. 19–26, Jan. 2020. doi: 10.1109/JSSC.2019.2939081

[10] J. F. Bulzacchelli, "Equalization for Electrical Links: Current Design Techniques and Future Directions," *IEEE Solid-State Circuits Mag.*, vol. 7, no. 4, pp. 23–31, Fall 2015. doi: 10.1109/MSSC.2015.2475996

[11] J. W. M. Bergmans, *Digital Baseband Transmission and Recording*. Springer, 1996, ch. Adaptive Reception, pp. 373–450.

[12] V. Balan, O. Oluwole, G. Kodani, C. Zhong, R. Dadi, A. Amin, A. Ragab, and M.-J. E. Lee, "A 15–22 Gbps Serial Link in 28 nm CMOS With Direct DFE," *IEEE J. Solid-State Circuits*, vol. 49, no. 12, pp. 3104–3115, 2014. doi: 10.1109/JSSC.2014.2349992

[13] H. Higashi et al., "A 5-6.4-Gb/s 12-channel transceiver with pre-emphasis and equalization," *IEEE J. Solid-State Circuits*, vol. 40, no. 4, pp. 978–985, Apr 2005. doi: 10.1109/JSSC.2005.845562

[14] S. Palermo, S. Hoyos, S. Cai, S. Kiran, and Y. Zhu, "Analog-to-Digital Converter-Based Serial Links: An Overview," *IEEE Solid-State Circuits Mag.*, vol. 10, no. 3, pp. 35–47, Aug. 2018. doi: 10.1109/MSSC.2018.2844603

[15] V. Stojanović et al., "Autonomous dual-mode (PAM2/4) serial link transceiver with adaptive equalization and data recovery," *IEEE J. Solid-State Circuits*, vol. 40, no. 4, pp. 1012–1026, Apr. 2005. doi: 10.1109/JSSC.2004.842863

[16] H.-J. Chi, J.-S. Lee, S.-H. Jeon, S.-J. Bae, Y.-S. Sohn, J.-Y. Sim, and H.-J. Park, "A Single-Loop SS-LMS Algorithm With Single-Ended Integrating DFE Receiver for Multi-Drop DRAM Interface," *IEEE J. Solid-State Circuits*, vol. 46, no. 9, pp. 2053–2063, Sep. 2011. doi: 10.1109/JSSC.2011.2136590

[17] J. Savoj et al., "Wideband flexible-reach techniques for a 0.5-16.3Gb/s fully-adaptive transceiver in 20nm CMOS," in *Proc. of the IEEE 2014 Custom Integrated Circuits Conf.* IEEE, Sep. 2014, pp. 1–4. doi: 10.1109/CICC.2014.6945980.

[18] S. Dasgupta, C.R. Johnson, and A.M. Baksho, "Sign-sign LMS convergence with independent stochastic inputs," *IEEE Trans. Inf. Theory*, vol. 36, no. 1, pp. 197–201, Jan. 1990. doi: 10.1109/18.50391

[19] D. Menin, A. De Pr, A. Bandiziol, W. Grollitsch, R. Nonis, and P. Palestri, "A Simple Simulation Approach for the Estimation of Convergence and Performance of Fully-Adaptive Equalization in High-Speed Serial Interfaces," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 9, no. 10, pp. 2079–2086, Oct. 2019. doi: 10.1109/TCPMT.2019.2911177

[20] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, no. 9, pp. 1349–1387, Sept 1985. doi: 10.1109/PROC.1985.13298

[21] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, no. 8, pp. 1151–1162, 1976. doi: 10.1109/PROC.1976.10286

[22] B. Casper, "Clocking Wireline Systems: An Overview of Wireline Design Techniques," *IEEE Solid-State Circuits Magazine*, vol. 7, no. 4, pp. 32–41, 2015. doi: 10.1109/MSSC.2015.2476015

[23] A. Sanders and J. DAmbrosia, "Designcon 2004 channel compliance testing utilizing novel statistical eye methodology," 2004.

[24] G. Balamurugan, B. Casper, J. Jaussi, M. Mansuri, F. O'Mahony, and J. Kennedy, "Modeling and Analysis of High-Speed I/O Links," *IEEE Trans. Adv. Packag.*, vol. 32, no. 2, pp. 237–247, May 2009. doi: 10.1109/TADVP.2008.2011366

[25] D. Oh, J. Ren, and S. Chang, "Hybrid Statistical Link Simulation Technique," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 1, no. 5, pp. 772–783, May 2011. doi: 10.1109/TCPMT.2011.2118209

[26] J. Ren and K. S. Oh, "Multiple Edge Responses for Fast and Accurate System Simulations," *IEEE Trans. Adv. Packag.*, vol. 31, no. 4, pp. 741–748, Nov. 2008. doi: 10.1109/TADVP.2008.2002201

[27] D. Menin, A. Bandiziol, W. Grollitsch, R. Nonis, and P. Palestri, "Design and Simulation of a 12 Gb/s Transceiver with 8-Tap FFE, Offset-Compensated Samplers and Fully-Adaptive 1-Tap Speculative/3-Tap DFE and Sampling Phase for MIPI A-PHY Applications," *IEEE Trans. Circuits Syst. II*, 2019. doi: 10.1109/TCSII.2019.2926152. [Online]. Available: https://ieeexplore.ieee.org/document/8752403/

[28] G. Signorini, C. Siviero, M. Telescu, and I. S. Stievano, "Present and future of I/O-buffer behavioral macromodels," *IEEE Electromagn. Compat.*, vol. 5, no. 3, pp. 79–85, 2016. doi: 10.1109/MEMC.0.7764256

[29] J. N. Tripathi, V. K. Sharma, and H. Shrimali, "A Review on Power Supply Induced Jitter," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 9, no. 3, pp. 511–524, Mar. 2019. doi: 10.1109/TCPMT.2018.2872608

[30] A. Cristofoli, P. Palestri, N. D. Dalt, and L. Selmi, "Efficient Statistical Simulation of Intersymbol Interference and Jitter in High-Speed Serial Interfaces," *IEEE Trans. Compon. Packag. Manuf. Technol.*, vol. 4, no. 3, pp. 480–489, Mar. 2014. doi: 10.1109/TCPMT.2013.2282530

[31] M. Dazzi, P. Palestri, D. Rossi, A. Bandiziol, I. Loi, D. Bellasi, and L. Benini, "Sub-mW multi-Gbps chip-to-chip communication Links for Ultra-Low Power IoT end-nodes," in *2018 IEEE Int. Symp. on Circuits and Syst. (ISCAS)*, May 2018. doi: 10.1109/ISCAS.2018.8351893 pp. 1–5.

[32] E. Denlinger, "Losses of Microstrip Lines," *IEEE Trans. Microw. Theory Techn.*, vol. 28, no. 6, pp. 513–522, Jun. 1980. doi: 10.1109/TMTT.1980.1130112

[33] W. Getsinger, "Microstrip Dispersion Model," *IEEE Trans. Microw. Theory Techn.*, vol. 21, no. 1, pp. 34–39, Jan. 1973. doi: 10.1109/TMTT.1973.1127911

[34] E. Hammerstad and O. Jensen, "Accurate Models for Microstrip Computer-Aided Design," in *MTT-S Int. Microwave Symp. Dig.*, vol. 80. MTT006, 1980. doi: 10.1109/MWSYM.1980.1124303 pp. 407–409.

[35] M. Kirschning and R. Jansen, "Accurate model for effective dielectric constant of microstrip with validity up to millimetre-wave frequencies," *Electron. Lett.*, vol. 18, no. 6, p. 272, 1982. doi: 10.1049/el:19820186

[36] H. Wheeler, "Transmission-Line Properties of a Strip on a Dielectric Sheet on a Plane," *IEEE Trans. Microw. Theory Techn.*, vol. 25, no. 8, pp. 631–647, Aug. 1977. doi: 10.1109/TMTT.1977.1129179

[37] M. Kirschning and R. Jansen, "Arguments and an accurate model for the power-current formulation of microstrip characteristic impedance," *Archiv der elektrischen Übertragung: AEÜ*, vol. 37, pp. 108–112, mar 1983.

[38] M. Kirschning and R. Jansen, "Accurate Wide-Range Design Equations for the Frequency-Dependent Characteristic of Parallel Coupled Microstrip Lines," *IEEE Trans. Microw. Theory Techn.*, vol. 32, no. 1, pp. 83–90, Jan. 1984. doi: 10.1109/TMTT.1984.1132616

[39] T. Brazil, "Causal-convolution – a new method for the transient analysis of linear systems at microwave frequencies," *IEEE Trans. Microw. Theory Techn.*, vol. 43, no. 2, pp. 315–323, Feb. 1995. doi: 10.1109/22.348090

[40] Alexander, J.D.H., "Clock recovery from random binary signals," *Electron. Lett.*, vol. 11, no. 22, pp. 541–542, Oct. 1975. doi: 10.1049/el:19750415

[41] V. Stojanović and M. Horowitz, "Modeling and analysis of high-speed links," in *Proc. of the IEEE 2003 Custom Integrated Circuits Conf., 2003.* IEEE, 2003, pp. 589–594. doi: 10.1109/CICC.2003.1249467.

[42] G. R. Gangasani et al., "A 32 Gb/s Backplane Transceiver With On-Chip AC-Coupling and Low Latency CDR in 32 nm SOI CMOS Technology," *IEEE J. Solid-State Circuits*, vol. 49, no. 11, pp. 2474–2489, Nov. 2014. doi: 10.1109/JSSC.2014.2340574

[43] J. Han, Y. Lu, N. Sutardja, K. Jung, and E. Alon, "A 60Gb/s 173mW receiver frontend in 65nm CMOS technology," in *2015 Symp. on VLSI Circuits (VLSI Circuits)*, June 2015, pp. C230–C231. doi: 10.1109/VLSIC.2015.7231268.

[44] T. Shibasaki et al., "A 56Gb/s NRZ-electrical 247mW/lane serial-link transceiver in 28nm CMOS," in *2016 IEEE Int. Solid-State Circuits Conf. (ISSCC)*. IEEE, Jan. 2016, pp. 64–65. doi: 10.1109/ISSCC.2016.7417908.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 2, 537-544 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Understanding Human Perception of Vibrotactile Feedback in Walking and Running Tasks

Emel Demircan*,[1], Elliot Recinos[2], I-Hung Khoo[3], Sharon Teng[4], Will Wu[5]

[1]*Mechanical and Aerospace Engineering, Biomedical Engineering, California State University Long Beach, 90803, U.S.A.*

[2]*Mechanical and Aerospace Engineering, California State University Long Beach, 90803, U.S.A.*

[3]*Electrical Engineering, Biomedical Engineering, California State University Long Beach, 90803, U.S.A.*

[4]*Department of Physical Therapy, California State University Long Beach, 90803, U.S.A.*

[5]*Department of Kinesiology, California State University Long Beach, 90803, U.S.A.*

ARTICLE INFO

Article history:
Received: 19 January, 2020
Accepted: 27 March, 2020
Online: 10 April, 2020

Keywords:
Haptic Feedback
Locomotion
Motion Training

ABSTRACT

Vibrotactile feedback is increasingly becoming an essential feedback component in several non-medical and medical areas. One area that vibrotactile feedback has not been explored as an intervention tool is in sports science. In addition, vibrotactile feedback lacks scientific evidence as a feedback mechanism within the sports world. A portable vibrotactile feedback system was developed to understand how a human perceives vibrotactile feedback while performing walking and running tasks. The system incorporates a model-based simulation framework to check human motion and visualize key performance metrics. We performed a pilot study on twenty-six subjects to understand the accuracy of human perception of vibrotactile feedback using the system we developed. Here, we investigated the type, number, and location of haptic feedbacks that would yield better subject perceived accuracy during walking and running. Findings suggested that staggered vibrotactile feedback would lead to higher accuracy than continuous vibrotactile feedback especially when more than two motors were used. For the subjects perception accuracy for all motor combinations, we observed an average that was 9.3% higher for walking compared to running. For one motor, the perception accuracy decreased by only 3.38% while running compared to walking. The decrease was much more significant for higher combinations of motors. The decrease in the perception accuracy of the motor locations was significantly less for both staggered ($p < 0.001$) and one motor ($p < 0.001$) vibration modes.

1 Introduction

This paper is an extension of work originally presented in IEEE International Conference on Ubiquitous Robots (URAI 2019) [1]. A key in quantifying and improving locomotion is identification of useful, tractable metrics and providing useful feedback. The most fundamental metric for our study, and arguably of most interest to many affected populations, is an estimate of injury risk. Despite the health benefits, a high number of lower extremity injuries can be linked to running [2, 3]. Estimates suggest that 10-20% of Americans run regularly, with 40-50% of these injured annually [4], although causation is more complex, with a survey of results across 17 published studies, involving a range specific population characteristics (age, experience, gender, etc.) showing annual injury rates can vary from 19% to 79% [5]. Among running injuries, 50% of injuries occur at the knee joint and the most common diagnosis is patellofemoral pain (PFP) [3, 5]. PFP can lead to severe pain and disability and is a precursor of knee osteoarthritis [6]. Real-time feedback is an effective method for motion and gait retraining. Augmented feedback has been shown to be a useful part in the process of learning a motor task and has been shown to enhance effects of walking and running gait programs [7, 8]. Different feedback modalities such as visual, audio, or vibrotactile feedback have been shown to effectively modify knee and impact loading [9]-[13]. Vibrotactile feedback is increasingly becoming an essential feedback component in several non-medical and medical areas [14]. Among

*Corresponding Author: Emel Demircan, 1250 Bellflower Blvd, ECS-645, California State University Long Beach, CA, 90840, emel.demircan@csulb.edu

vibrotactile feedback, coin vibration motors are chosen for being non-invasive, safe, low cost, small size, and its ability to be effective when placed at almost anywhere on the body [15]-[18]. Vibrotactile feedback can also be a means of delivering cues to a user learning new motor skills [19] or for patients undergoing rehabilitation therapy [20]. Models that are data-driven have been implemented to train gaits involving a combination of different kinematic modifications [21]. One study compared how combining haptic feedback with visual feedback had a greater effect in the learning performance for lower extremity tasks than the feedbacks alone [22]. A different study pointed out that wearable feedbacks in the form of haptic and auditory are commonly used in clinical applications for human gait [23]. Knee and trunk motion are the gait parameters most associated with retraining using wearable feedback [23]. Both visual-auditory feedback and visual-tactile feedback provide advantages in reducing reaction times and improving performance [24]. It has been observed that combining visual feedback with tactile feedback works best when performing multiple tasks that have high cognitive workload conditions [24]. In a study that evaluated assistive navigation systems for the blind, auditory feedback resulted in a 22 times higher cognitive load than haptic feedback [25]. Previous studies have observed that vision feedback provides a high degree of precision [12]. Vibration provides simple and intuitive feedback, particularly when vision is otherwise occupied [21]. In addition, vibration conveys Cartesian space directional cues well. Another benefit of haptic feedback systems is that they can provide feedback to multiple body segments and thus allow modifications of multiple gait parameters [21]. Motor tasks can be guided by using immediate feedback on complex tasks to provide proprioceptive information [14]. Several studies have shown how effective feedback can be in gait retraining and learning a task that can also slow down the progression of joint diseases [9]-[11]. Research still does not provide insight on how feedback needs to be provided and combined in the system to yield better results. It is important to maintain latency as low as possible since it is highly dependent on the motor task and type of vibrotactile stimuli.

To enhance motor learning, visual feedback can be provided in the process of gait retraining. This type of feedback facilitates external focus which has been shown to be beneficial for motor learning due to its ability of promoting automaticity in movement control [26]. Depending on the timing that the feedback is delivered, it can be divided into concurrent (online, real-time) and terminal (at the end) [27]. Concurrent feedback has been reported to enhance learning of complex motor tasks [28] and is especially helpful in early phase of learning [27]. Knowledge of Performance and Knowledge of Results are the two categories of Terminal feedback. While knowledge of performance provides feedback regarding the execution of the task (i.e. joint kinematics and kinetics of gait), knowledge of results pertains information related to the response outcome rather than about the movement (i.e. VO2 consumption during gait) [29]. Terminal feedback has been found to promote motor learning and facilitates motor retention [27]. Despite studies that focus on multi-modal feedback mechanisms on human kinematics, there is a lack of understanding on how feedback interfaces can be used with cyber-physical systems to improve response and engagement. Several studies demonstrate that there is not enough evidence yet to support implementation of tactile feedback to enhance performance [7, 16, 30].

In addition to augmented feedback systems, there has been much research in developing smart wear with integrated sensors, with applications in different areas. Some recent examples that involve soft suits which can be used in daily activities include the following: (i) A sensing suit using elastic tights and hyper-elastic strain sensors was developed for lower limb biomechanics measurements [31]; (ii) A knee pad and anklet wearable system that contains soft stretchable sensors was developed to track and analyze the joint position of lower limbs [32]; (iii) An IMU integrated shirt, with conductive yarn for circuit patterning and flexible printed circuit boards for interconnections, was designed for posture monitoring [33]; (iv) Leggings with stretch sensors were designed to detect knee valgus [34], commonly observed in athletes at risk of knee injuries; (v) A soft exo-suit, which consists of: a waist belt, leg strap, calf wrap, insole, load cells, gyroscopes, and Bowden cables, was developed to help improve walking in post-stroke patients [35]; (vi) The commercial product by Athos includes a compression shirt and shorts with integrated IMU and EMG sensors, and provides visual information on muscle efforts using a smartphone app (www.liveathos.com); (vii) a portable gait asymmetry rehabilitation system that delivers vibrotactile cues based on gait phase measurement to improve gait symmetry for individuals with stroke [20, 36].

It remains unclear how haptic mode (e.g., simultaneous or staggered) or number of motors affects individuals perception of haptic feedback during locomotion. In this study, we present a portable haptic feedback system that uses actuators [37]-[39] to provide the wearer with real-time haptic feedback during locomotion. We designed the system with the wearers comfort during locomotion in mind by making it portable and lightweight. Using this system, we conducted a pilot study to examine subject perceived haptic feedback during different types of haptic mode and numbers of motors during walking and running. Coin-based vibrotactile actuators were chosen and mounted on custom 3D printed cases to avoid impeding the human bodys natural motion. We had 3 main questions in mind when we designed the experiments: (i) How many vibration motors can a subject perceive at once?; (ii) Where on the subjects body should the haptic feedback be provided?; (iii) How effective are the subjects when interpreting higher orders of haptic feedback occurring simultaneously?

Our study is an extension of work originally presented in IEEE International Conference on Ubiquitous Robots (URAI 2019) [1]. Here, we provide a more comprehensive understanding of human interpretation of vibrotactile feedback during walking and running, including the effect of feedback locations. For future work, we plan to train subjects on how interpret haptic cues to increase their performance using this framework and suit by altering key kinematic variables during locomotion.

2 Vibrotactile Feedback

For the preliminary experiment, we developed a vibrotactile feedback system that consists of the following three subsystems: (1) the hardware, that is comprised of 10 vibrotactile motors (10 x 2.7 mm Coin Mobile Phone Vibration Motor) and 1 Bluetooth-enabled Arduino-Nano microcontroller, (2) an in-house made GUI that sends

Bluetooth commands to the vibrotactile systems microcontroller, and (3) a custom button panel that is placed on a treadmill to record the subjects haptic response (Figure 1).

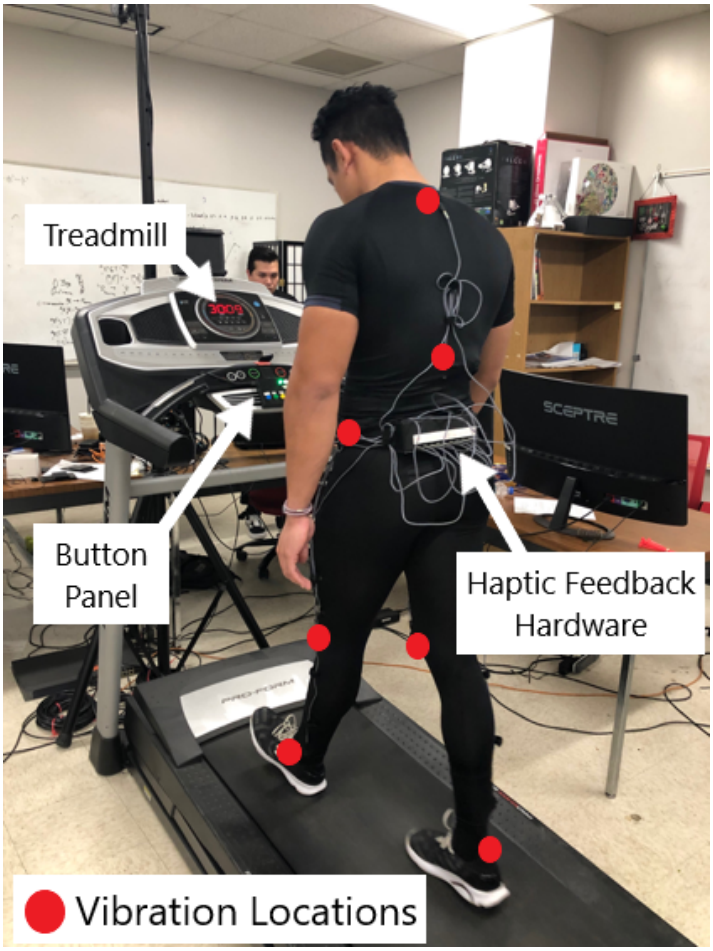


Figure 1: Subject wearing the wireless sensing and feedback system.

Our frameworks GUI is a Windows Presentation Foundation (WPF) application, that uses the 32feet.NET library to discover and connect to the vibrotactile feedback systems Bluetooth module, and a wired serial communication protocol to connect to the custom button panel. When connected to the haptic suit system, the experimenter can use the GUI to manually actuate the 200Hz (slightly below the peak sensitivity of fast-acting mechanoreceptors [40]) haptic suit motors and motor combinations by clicking on the appropriate radio buttons. Figure 2 shows how the motor actuation type (continuous or staggered) and duration of actuation can be manually changed by the experimenter using the GUI. Note that the GUI also provides visual feedback to the experimenter on what motors are being actuated by coloring in the circles that correspond to activated motors on the GUIs body model. When the Start Auto button is pressed, the GUI performs an automatic procedure in which motor actuation commands are executed. Specifically, the sequence and order of actuation modes are based on the selected Subject Number while the time gap between each actuation is randomly chosen at run time to be the subject-specific sequence.

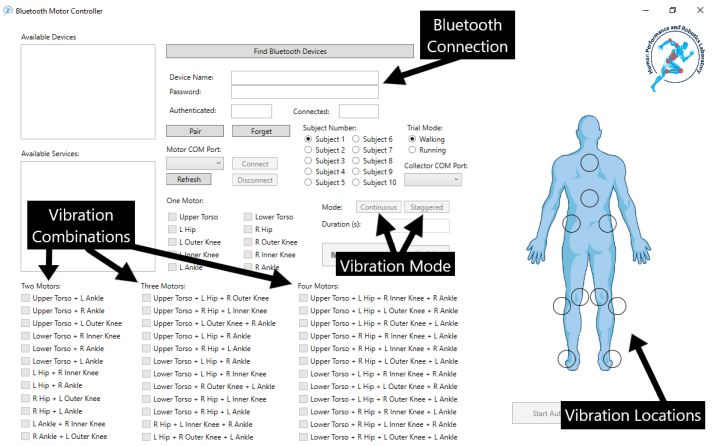


Figure 2: GUI for the vibrotactile feedback experiment.

While the automatic procedure is running, a separate thread is instantiated and used to collect the subject response data outputted by the custom button panel. The GUI automatically logs: the haptic command events, the subjects response input, and the timestamps at which these events occur, during the automatic procedure of the main process thread. When the automatic procedure finishes, the haptic events and subject response events are combined with the timestamps and then analyzed based on each events unique keywords. The subject response events are grouped with their respective actuator commands (by determining the closest actuator command event to the response event) and, with the use of specific event keywords, compared with their respective actuator commands to quantify the accuracy of the subjects response. For example, the subjects response will be reported as false (zero) if the subjects response is TORSO when the haptic feedback command was activated as LOWER TORSO + L KNEE (Continuous) since the subject failed to perceive the correct feedback. The software then calculates statistical results with the use of the ClosedXML library and saves both statistical and experimental data as an Excel file.

3 Simulation Framework

The simulation framework was developed to read existing OpenSim models and build them with the Unity game engine environment. OpenSims musculoskeletal properties and definitions are implemented by the framework. This allows the framework to acquire unique behaviors that are only found in OpenSims software. A few unique behaviors include muscle path joints that move during motion, conditional muscle path joints, and application of cubic spline for joint connectivity. The framework takes advantage of Unitys editor window to allow the user, while in edit mode, to upload and adjust dynamic OpenSim models. The simulation framework is capable of symbolic calculus and matrix operations due to implementations of the ALGIB numerical analysis and data-processing library [41], C++ Mathematical Expression Toolkit Library (ExprTK) [42], MITs Math.NET Symbolics [43]. At runtime of the framework, the user can symbolically solve both kinematic and muscular Jacobians of an OpenSim multi-body model using the Math.NET Symbolics library. This allows the user to generate the

symbolic expression that can used by the ExprTk library. Due to the high computation expense of symbolic interpretations, dynamic compilation functionality was implemented. Dynamic compilation allows the framework to compile and save symbolic expressions to be used in future runtimes. The simulation framework has the capability to solve complex and dynamic computations during motion tracking due to the implementation of dynamic compilation with Math.Net Symbolics closed-form symbolic computation. The framework can perform these computations at a framerate of 100 FPS or higher depending on the computers specifications. The framework generates generalized coordinates by implementing algorithms to decompose motion of a body [44]-[46] and can act as an interface for generic motion tracking.

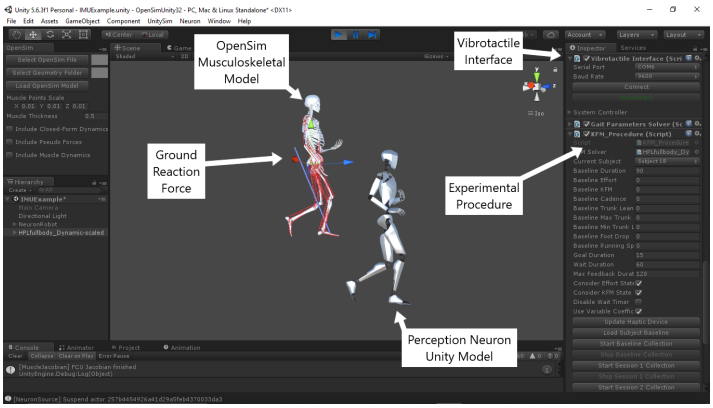


Figure 3: Simulation Framework.

To evaluate the motion tracking interfaces performance, a Perception Neuron full-body IMU suit was integrated into the framework. As expected, Noitom already provides the software (Axis Neuron) and the rigged humanoid model required to use the IMU motion capture system within Unity. As designed, the only task that must be completed to use the IMU motion capture system with our framework is to connect the individual limbs of the humanoid model to the appropriate JointTracker elements. The ground reaction forces are estimated by our framework using the motion data from IMU and the subject-specific mass matrix. The simulation framework provides haptic feedback commands to an arduino nano microcontroller that is bluetooth enabled. The simulation framework connects to the microcontroller through a haptic feedback interface script that connects to the Bluetooths communication port at 9600 baud rate (Figure 3). Once connected, the simulation runs an experimental procedure script that tracks different kinematic metrics and sends vibration commands in the form of strings to the microcontroller.

Effective indicators of joint loading that can be used in injury prevention and sports performance are joint moments. Progression of patellofemoral joint (PFJ) osteoarthritis can be related to flexion moment impulse and peak knee flexion [47]. High PFJ reaction force and stress has been associated to increased knee flexion moment that suggests requirements for greater quadriceps force [48, 49]. Our framework estimates the 3-dimensional joint moments in real-time using the inverse dynamics algorithm. Both the kinematic and kinetic values estimated by our framework were validated against the data reported in the literature [47]. The novel simulation framework and the motion/force data will be used in our future study to estimate joint contact forces and other performance metrics (i.e. metabolic cost) during locomotion.

4 Experiments

This section presents the experimental evaluation, including experimental setup, subjects and data collection, and statistical analysis.

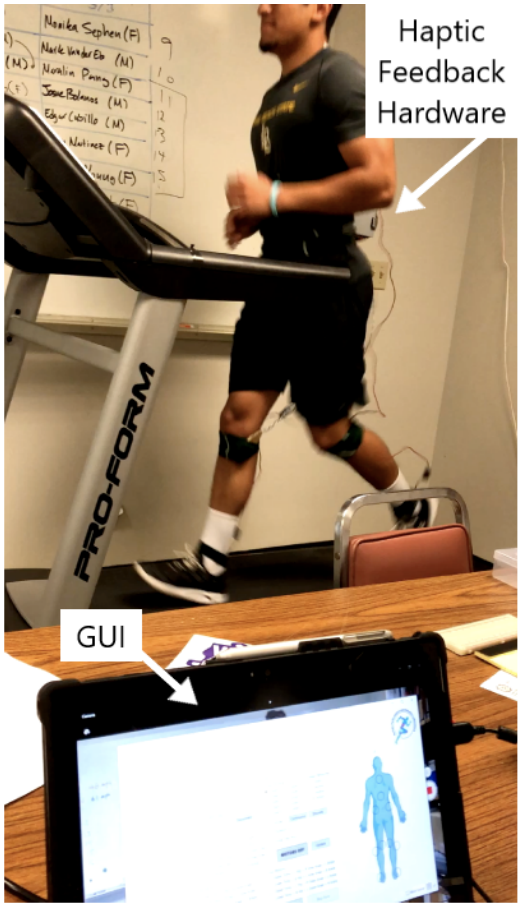


Figure 4: Experimental Setup and System GUI.

4.1 Subjects

Eighteen subjects (10 males, 8 females, age: 22.5 $\pm$ 3.0 years, BMI: 22.6 $\pm$ 3.6 kg/m^2) participated in the study. All subjects reported no pain or symptoms while walking or running, nor history of surgery or sensory deficit. To have scientifically correct subject-specific scaling for the modeling purposes in our future work, we did not include children or the elderly. Inclusion criteria for subject recruitment are the following: (1) between ages of 18 and 40 years old; (2) familiar with running on a treadmill; and (3) run at least 8 miles/week for 4 weeks prior to participation. Exclusion criteria for subject recruitment are the following: (1)History of lower extremity or low back surgery that may affect running kinematics, muscle activation, or kinetics; (2) low back or lower extremity discomfort or pain during the data collection or within 3 months prior of the experiment; and (3) any mental or physical condition that prevents

the subject from performing the exercises or tasks safely. Medical clearance and written informed consent forms approved by the California State University, Long Beach Human Subjects Review Board (IRB : 1300259-2) were obtained from the participants.

4.2 Experimental Setup

We used the same experimental setup detailed in our previous work [1]. The experiment consisted of 4 trials that got divided into walking and running. Subjects were tasked to perform two walking and two running trials on a treadmill at self-selected comfortable speed (walking: 1.1 ± 0.2 m/s, running: 2.2 ± 0.2 m/s). Each trial lasted between 12 to 14 minutes. During each trial, subjects experienced vibrations by motors placed over 10 locations of his/her body bony landmarks: upper and lower back, bilateral posterior superior iliac spines, bilateral medial and lateral femoral epicondyles, and bilateral lateral malleoli.

Two types of vibrations were implemented, continuous and staggered. During continuous vibration, the combinations of motors vibrated at the same time. During staggered vibrations, the motors vibrated in a sequential manner for the combination of motors. A 0.5 second delay was programmed between each motor during the vibration sequence. The vibration commands included up to a combination of four motors in a pre-determined random order. Subjects were tasked to verbally report the number and location of the vibration motors they perceived during the experiment. The subject was also tasked to report the vibration locations using the button panel on the treadmill. Subjects were informed if the trial was continuous or staggered with verbal instructions and an LED light on the button panel.

4.3 Data Collection

Feedback perception accuracy (FPA), which is defined as the percent accuracy of detecting correct number and location of vibration motors, was calculated for each combination. The average reaction time for each subject was collected during each trial. Subjects comfort level of wearing the system during locomotion was immediately reported after each experiment using a ten-point bipolar Likert-type scale.

4.4 Statistical Analysis

Two- way repeated measures ANOVA was used to evaluate effects of number of motors (i.e., 2, 3, and 4) and vibration modes (i.e., continuous and staggered) on FPA during walking and running. The FPA (%) between the first and second trial for walking and running and for the different vibration modes (staggered or continuous) was compared using paired-sample t-tests. Separate t-tests were performed independently for male (M) and female (F) subjects. An alpha of 0.05 was set for the statistical significance level.

5 Results

Overall, FPAs were greater than 50% in all motor and activity conditions in staggered mode. On the other hand, FPAs reached above 50% only during two motor condition in continuous mode. Results of ANOVAs showed significant number-of-motor, vibration-mode, and interaction effects ($p < 0.001$) for both walking and running. FPAs reduced as the number of motors increased. Compared to continuous mode, staggered mode yields higher FPAs, especially when the number of motors increased. When examining staggered or continuous mode alone, FPAs were significantly higher during walking than running ($p = 0.04$ and 0.01, respectively) with a mean difference of 8%.

Table 1: Table P-Values, averages, and standard deviations for walking and running trials

Task	**p-Value**	**Trial 1** (%)	**SD** (%)	**Trial 2** (%)	**SD** (%)
Walking	0.584	60.63	11.50	61.99	14.09
Running	0.384	54.40	12.64	56.10	12.07

Table 1 shows the respective FPA p-values, averages, and standard deviations for trial 1 and trial 2 during walking and running. The FPA means for walking and running, respectively, showed an increase of 2.23% and 3.12% that suggests effective learning for both types of locomotion.

Table 2: Table P-Values, averages, and standard deviations for males and females

Task	**p-Value**	**Male** (%)	**SD** (%)	**Female** (%)	**SD** (%)
Walking	0.052	66.09	10.74	55.33	10.74
Running	0.136	58.96	9.39	50.61	13.18

Table 2 shows the respective FPA p-values, averages, and standard deviations for males and females during walking and running. A decrease of 12.60% in the standard deviation from walking to running was observed among the male population. An increase of 22.70% in the standard deviation from walking to running was observed among the female population.

Table 3: Table P-Values, averages, and standard deviations for vibration modes

Task	**p-Value**	**S** (%)	**SD** (%)	**C** (%)	**SD** (%)
Walking	0.0001	50.72	9.14	74.84	17.24
Running	0.0001	44.14	8.23	69.05	18.54

Table 3 shows the respective FPA p-values, averages, and standard deviations for continuous (C) vibrations and staggered (S) vibrations during walking and running. A decrease of 9.95% in the standard deviation from walking to running was observed within the continuous vibration mode. An increase of 7.49% in the standard deviation from walking to running was observed within the staggered vibration mode. Figure 6 shows the FPA averages for 1, 2, 3, and 4 motors. Most notably, 1 motor yielded an FPA of 92.78 for walking.

The overall FPA decreased when the subjects were asked to verbally identify the location of the vibrotactile devices they felt. The average decrease for detecting one motor location was 2.98% during

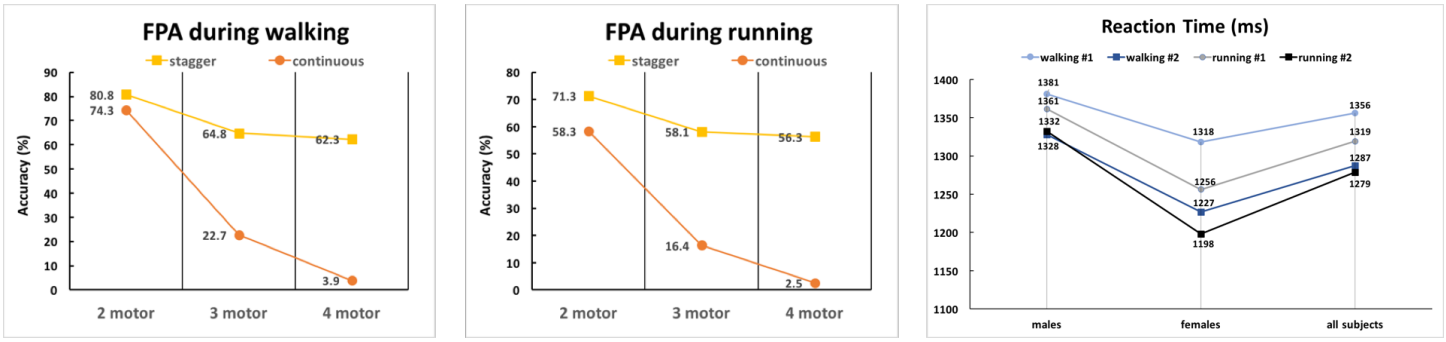


Figure 5: FPA during walking (left), FPA during running (middle), and reaction time (right).

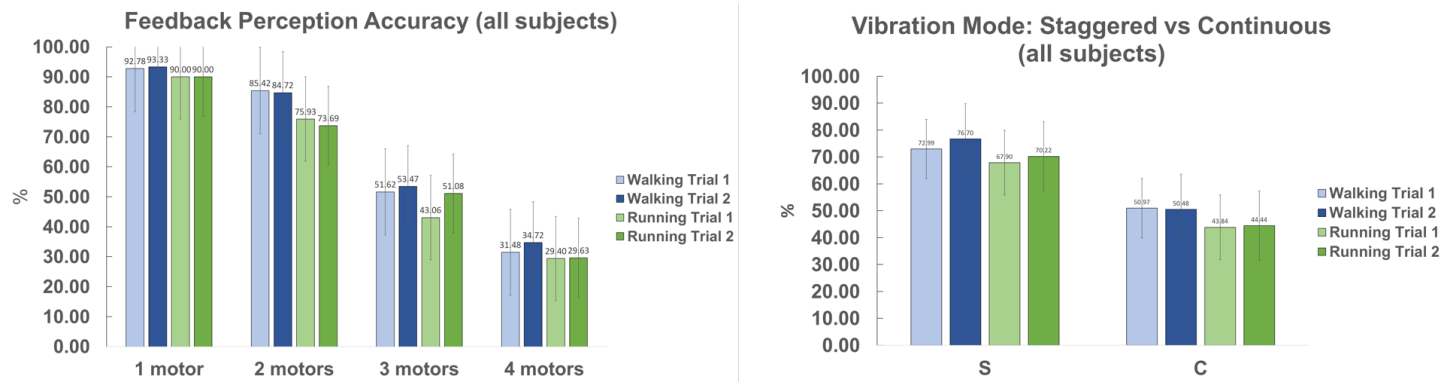


Figure 6: FPA for subjects (left); Vibration modes: staggered vs. continuous (right).

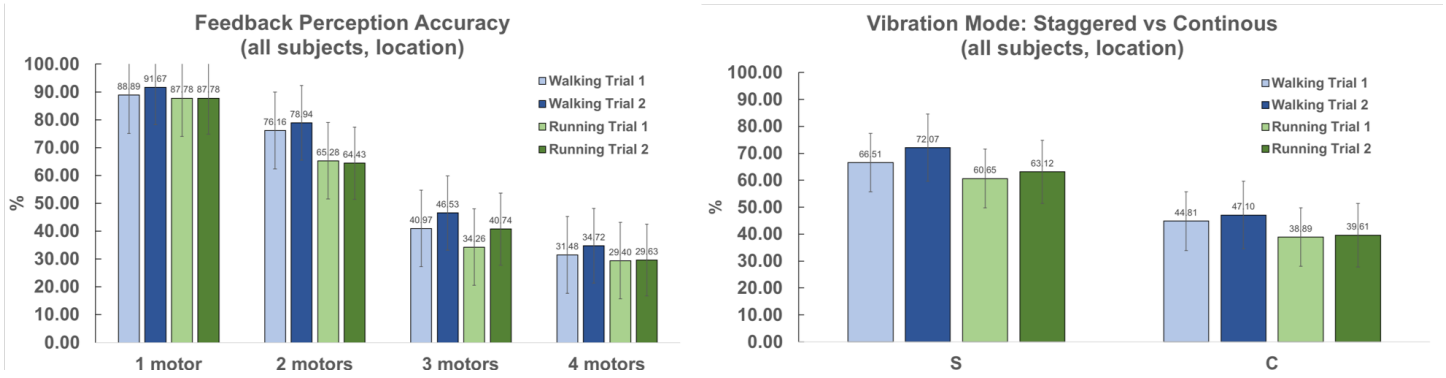


Figure 7: FPA of the feedback locations on subjects (left); Vibration modes: staggered vs. continuous (right).

walking and 2.46% during running (see Figure 7 (left)). For staggered vibration mode, there was an average of 7.45% and 10.39% decrease in FPA of the correct location of vibrotactile motors during walking and running, respectively. For continuous vibration mode, the decrease was more significant: 9.39% and 11.08% decrease in FPA of the correct location of vibrotactile motors during walking and running, respectively (see Figure 7 (right)).

The results from Figure 5 (left and middle) suggest that staggered vibrations are easier to perceive by humans. This can be observed by comparing the loss in FPA as the number of simultaneous stimuli increases. Figure 5 (left and middle) show a greater loss (slope) for continuous vibrations as the number of motors is increased compared to staggered, which has less of a negative slope. This may suggest that cognitive load is less when trying to identify vibrations that are staggered throughout body locations compared to simultaneous continuous vibrations. This shows promise when trying to maintain low cognitive load for motor and gait training. Figure 6 (right) and Figure 7 (right) both reinforce that staggered is much more suitable when you want a person to perceive total number of stimulation and accurately distinguish among different locations. Another possible reason that staggered is a more suitable form of haptic feedback for motor and gait training is because it has been observed that continuous vibration on the same location may lead to numbness of that section of the skin [50]. Meanwhile, staggered vibrations provide the user with rest periods that may reduce the numbness associated to constant stimulation. As mentioned in the experiment section, subjects reported an average comfort level score of 8.4 using the bipolar Likert-type with a ten-point scale. Subjects average reaction times can be observed in Figure 5 (right).

6 Discussion

For this paper, we explored human perception of vibrotactile feedback using a portable system that provides vibrations during walking and running tasks. Findings from our previous study gave insight of how vibrotactile feedback is perceived by humans during locomotion. This study showed that staggered vibrotactile feedback would yield greater subject perceived accuracy during locomotion especially when more than two motors were used. This study provides valuable insights into the feasibility of vibration mode and number of motors during locomotion. Information obtained from this study can be used to improve the haptic feedback systems design and experimental protocol.

This studys results set a foundation to develop multi-modal feedback systems to improve performance during walking and running. We plan to incorporate the simulation framework into the multi-modal system to provide different feedback modalities for coordinated motions. Previous studies have observed that vision feedback provides a high degree of precision [12]. Vibration provides simple and intuitive feedback, particularly when vision is otherwise occupied [21]. In addition, vibration conveys Cartesian space directional cues well. By integrating visual, audio, and haptic feedback, we will combine their advantages to provide rich and intuitive feedback to improve human performance and decrease the risk of injuries during locomotion.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment The design of the study and collection, analysis, and interpretation of data were funded by California State University Long Beach Faculty Seed Grants.

References

[1] Demircan, E., Recinos, E., Abella, J.R., Khoo, I., Teng, S., Wu, W.: Perception accuracy of vibrotactile feedback during locomotion. In: 16th International Conference on Ubiquitous Robots, UR 2019, Jeju, South Korea, June 24-27, 2019, pp. 673–677 (2019). doi:10.1109/URAI.2019.8768771

[2] Koplan, J.P., Rothenberg, R., Jones, E.: The natural history of exercise: a 10-yr follow-up of a cohort of runners. Medicine & Science in Sports & Exercise **27**(8), 1180–1184 (1995)

[3] Taunton, J.E., Ryan, M.B., Clement, D.B., McKenzie, D.C., Lloyd-Smith, D.R., Zumbo, B.D.: A retrospective case-control analysis of 2002 running injuries. British journal of sports medicine **36**(2), 95–101 (2002)

[4] Fields, J.C..W.K.M..J.J.C. Karl B.1; Sykes: Prevention of running injuries. Current Sports Medicine Reports **9**, 176–182 (2010)

[5] van Gent, R.N., Siem, D., van Middelkoop, M., van Os, A.G., Bierma-Zeinstra, S.M.A., Koes, B.W., Taunton, J.E.: Incidence and determinants of lower extremity running injuries in long distance runners: a systematic review. British journal of sports medicine **41**(8), 469–480 (2007)

[6] Thomas, M.J., Wood, L., Selfe, J., Peat, G.: Anterior knee pain in younger adults as a precursor to subsequent patellofemoral osteoarthritis: a systematic review. BMC Musculoskeletal Disorders **11**(1), 201 (2010)

[7] Shull, P.B., Damian, D.D.: Haptic wearables as sensory replacement, sensory augmentation and trainera review. J Neuroeng Rehabil **12**(1) (2015)

[8] Agresta, C., Brown, A.: Gait Retraining for Injured and Healthy Runners Using Augmented Feedback: A Systematic Literature Review. Journal of Orthopaedic and Sports Physical Therapy **45**(8), 576–584 (2015)

[9] Barrios, J.A., Crossley, K.M., Davis, I.S.: Gait retraining to reduce the knee adduction moment through real-time visual feedback of dynamic knee alignment. J Biomech Eng. **43**(11), 2208–2213 (2012). doi:10.1016/j.jbiomech.2010.03.040

[10] Riskowski, J.L., Mikesky, A.E., Bahamonde, R.E., Burr, D.B.: Design and validation of a knee brace with feedback to reduce the rate of loading. J Biomech Eng. **131**(8) (2009). doi:10.1115/1.3148858

[11] Dowling, A.V., Fisher, D.S., Andriacchi, T.P.: Gait modification via verbal instruction and an active feedback system to reduce peak knee adduction moment. J Biomech Eng. **132**(7) (2009). doi:10.1115/1.4001584

[12] Wheeler, J.W., Shull, P.B., Besier, T.F.: Real-time knee adduction moment feedback for gait retraining through visual and tactile displays. J Biomech Eng. **133**(4) (2011). doi:10.1115/1.4003621

[13] Crowell, H.P., Milner, C.E., Hamill, J., Davis, I.S.: Reducing Impact Loading During Running with the Use of Real-Time Visual Feedback. J of Orth and Sports Phy Ther **40**(4), 206–213 (2010)

[14] VanBreda, E., Verwulgen, S., Saeys, W., Wuyts, K., Peeters, T.: Vibrotactile Feedback as a Tool to Improve Motor Learning and Sports Performance: A Systematic Review. BMJ Open Sport and Exercise Medicine **3**, 21–53 (2017)

[15] Bark, K., Hyman, E., Tan, F., Cha, E., Jax, S.A., Buxbaum, L.J., Kuchenbecker, K.J.: Effects of Vibrotactile Feedback on Human Learning of Arm Motions (V), 1–42 (2015). doi:10.1109/TNSRE.2014.2327229.Effects

[16] Alahakone, A.U., Senanayake, S., Arosha, M.: Vibrotactile feedback systems: current trends in rehabilitation, sports and information display. In: Advanced Intelligent Mechatronics (2009). ASME International Conference

[17] Bannwart, M., Pyk, P., Kiper, D., Eng, K., Gassert, R., Kim, Y.: 2013 International Conference on Virtual Rehabilitation : Philadelphia, Pennsylvania, USA, 26-29 August 2013.

[18] Weber, P., Ruekert, E., Calandra, R., Peters, J., Beckerle, P.: IEEE RO-MAN 2016 : the 25th IEEE International Symposium on Robot and Human Interactive Communication : 25th Anniversary : August 26 to August 31, 2016, Teachers College, Columbia University, New York, U.S.A.

[19] Van Der Linden, J., Schoonderwaldt, E., Bird, J., Johnson, R.: MusicJacket - Combining motion capture and vibrotactile feedback to teach violin bowing. IEEE Transactions on Instrumentation and Measurement **60**(1), 104–113 (2011). doi:10.1109/TIM.2010.2065770

[20] Afzal, M.R., Lee, H., Yoon, J., Oh, M.K., Lee, C.H.: Development of an augmented feedback system for training of gait improvement using vibrotactile cues. 2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence, URAI 2017, 818–823 (2017). doi:10.1109/URAI.2017.7992833

[21] Shull, P.B., Lurie, K.L., Cutkosky, M.R., Besier, T.F.: Training multi-parameter gaits to reduce the knee adduction moment with data-driven models and haptic feedback. J Biomech Eng. **44**(8), 1605–1609 (2011). doi:10.1016/j.jbiomech.2011.03.016

[22] Koritnik, T., Koenig, A., Bajd, T., Riener, R., Munih, M.: Comparison of visual and haptic feedback during training of lower extremities. Gait and Posture **32**(4), 540–546 (2010). doi:10.1016/j.gaitpost.2010.07.017

[23] Shull, P.B., Jirattigalachote, W., Hunt, M.A., Cutkosky, M.R., Delp, S.L.: Quantified self and human movement: A review on the clinical impact of wearable sensing and feedback for gait analysis and intervention. Gait and Posture **40**(1), 11–19 (2014). doi:10.1016/j.gaitpost.2014.03.189

[24] Burke, J.L., Prewett, M.S., Gray, A.A., Yang, L., Stilson, F.R.B., Coovert, M.D., Elliot, L.R., Redden, E.: Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. In: Proc. of the 8th International Conference on Multimodal Interfaces, Canada, pp. 108–117 (2006)

[25] Martinez, M., Constanttinescu, A., Schauerte, B., Koester, D., Steifelhagen, R.: Cognitive evaluation of haptic and audio feedback in short range navigation tasks. In: Proc. of the International Conference on Computers for Handicapped Persons, ICCHP, pp. 128–135 (2014)

[26] Shea, C.H., Wulf, G.: Enhancing motor learning through external-focus instructions and feedback. Human movement science **18**, 553–571 (1999)

[27] Sigrist, R., Rauter, G., Riener, R., Wolf, P.: Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review. Psychonomic Bulletin & Review **20**(1), 21–53 (2012)

[28] Wulf, G., Shea, C.H.: Principles derived from the study of simple skills do not generalize to complex skill learning. Psychonomic Bulletin & Review **9**(2), 185–211 (2002)

[29] Salmoni, A.W., Schmidt, R.A., Walter, C.B.: Knowledge of results and motor learning: a review and critical reappraisal. Psychological bulletin **95**(3), 355–386 (1984)

[30] Choi, S., Kuchenbecker, K.J.: Vibrotactile display: Perception, technology, and applications. In: Proceedings of the IEEE, pp. 2093–2104 (2013). IEEE

[31] Mengüç, Y., Park, Y.-L., Martinez-Villalpando, E., Aubin, P., Zisook, M., Stirling, L., Wood, R.J., Walsh, C.J.: Soft wearable motion sensing suit for lower limb biomechanics measurements. In: Robotics and Automation (ICRA), 2013 IEEE International Conference On, pp. 5309–5316 (2013). IEEE

[32] Totaro, M., Poliero, T., Mondini, A., Lucarotti, C., Cairoli, G., Ortiz, J., Beccai, L.: Soft smart garments for lower limb joint position analysis. Sensors **17**(10), 2314 (2017)

[33] Kang, S.-W., Choi, H., Park, H.-I., Choi, B.-G., Im, H., Shin, D., Jung, Y.-G., Lee, J.-Y., Park, H.-W., Park, S., *et al.*: The development of an imu integrated clothes for postural monitoring using conductive yarn and interconnecting technology. Sensors **17**(11), 2560 (2017)

[34] Pettys-Baker, R., Compton, C., Utset-Ward, S., Tompkins, M., Holschuh, B., Dunne, L.E.: Design and development of valgus-sensing leggings. In: 2017 Design of Medical Devices Conference, pp. 001–0501700105017 (2017). American Society of Mechanical Engineers

[35] Awad, L.N., Bae, J., Odonnell, K., De Rossi, S.M., Hendron, K., Sloot, L.H., Kudzia, P., Allen, S., Holt, K.G., Ellis, T.D., *et al.*: A soft robotic exosuit improves walking in patients after stroke. Science translational medicine **9**(400), 9084 (2017)

[36] Afzal, M.R., Lee, H.-s., Oh, M.-k., Lee, J.-h., Yoon, J.: Effects of Vibrotactile Feedback on Human Learning of Arm Motions. BioMed Research International **2015**, 1–4 (2014)

[37] Rogers, J.A.: A clear advance in soft actuators. Science **341**, 968–969 (2013)

[38] Majidi, C.: Soft robotics: A perspective - current trends and prospects for the future. Soft Robot **1**, 5–11 (2014)

[39] Yeo, W.-H., Kim, Y.-S., Lee, J., Ameen, A., Shi, L., Li, M.: Multifunctional epidermal electronics printed directly onto the skin. Adv Mater **25**, 2773–2778 (2013)

[40] Verrillo, R.T.: Age related changes in the sensitivity to vibration. Journal of gerontology **35**(2), 185–193 (1980)

[41] C++/C# Numerical Analysis Library

[42] C++ Mathematical Expression Library (ExprTk)

[43] Math.NET Symbolics

[44] Mladenova, C., Mladenov, I.: Vector decomposition of finite rotations. Reports on Mathematical Physics (2011)

[45] Brezov, D., Mladenova, C., Mladenov, I.: Vector decomposition of rtations. Geometry and Symmetry in Physics (2012)

[46] Dobrowolski, P.: Swing-twist decomposition in clifford algebra. CoRR (2015)

[47] Teng, H.-L., MacLeod, T.D., Link, T.M., Majumdar, S., Souza, R.B.: Higher knee flexion moment during the second half of the stance phase of gait is associated with the progression of osteoarthritis of the patellofemoral joint on magnetic resonance imaging. Journal of Orthopaedic and Sports Physical Therapy **45**(9), 656–664 (2015)

[48] Ho, K., Blanchette, M., CM., P.: The influence of heel height on patellofemoral joint kinetics during walking. Current Sports Medicine Reports **35**, 271–275 (2012)

[49] Teng, H.-L., C.M., P.: Sagittal plane trunk posture influences patellofemoral joint stress during running. J Orthop Sports Phys Ther. **44**(10), 785–792 (2014)

[50] Guo, W., Ni, W., Chen, I.M., Ding, Z.Q., Yeo, S.H.: Intuitive vibro-tactile feedback for human body movement guidance. 2009 IEEE International Conference on Robotics and Biomimetics, ROBIO 2009, 135–140 (2009). doi:10.1109/ROBIO.2009.5420612

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 5, No. 2, 567-577 (2020)
www.astesj.com
Special Issue on Multidisciplinary Sciences and Engineering

ASTES Journal
ISSN: 2415-6698

Securing Hybrid SDN-based Geographic Routing Protocol using a Distributed Trust Model

Lylia Alouache[*,1], Mohamed Maachaoui[2], Rachid Chelouah[1]

[1] *ETIS CNRS ENSEA UMR 8051. Computer science, CY Cergy Paris University, 95000, France*

[2] *Quartz Laboratory. Computer science, CY Cergy Paris University, 95000, France*

ARTICLE INFO

Article history:
Received: 15 January, 2020
Accepted: 10 March, 2020
Online: 12 April, 2020

Keywords:
IoV
Routing protocols
SDN
QoS
Security
Availability
Reliability
Integrity
Trust

ABSTRACT

In this paper, the vulnerabilities and the security attacks against vehicular networks, SDN architecture and some security solutions for SDVN are studied. Secondly, a complete and improved version of the secure HSDN-GRA routing protocol based on a distributed trust model is proposed where public keys of the vehicles are managed in a distributed way. Besides, in this approach, a weight is assigned to every vehicle, which is calculated from its freeload and its trust values per those of the neighbors. The trust value is deduced from the historical interactions stored in a log of communication errors. To measure the trust value, three unreliable behaviors occurred in a time interval are considered. The vehicle with the role of cluster head hosts a list of each vehicles misbehaviors forming a log of communication errors. A vehicle will be chosen as the next relay according to its weight. As a positive result of this complete proposed approach, the following security requirements are achieved: the vehicle's authentication and the data integrity are guaranteed by a signature mechanism, whereas an encrypted function is used to ensure the confidentiality of the exchanged data. The goal is to protect the routing process against malicious and unstable nodes. Finally, the implementation details and simulation analysis are given, and a comparative study between the secure and the insecure HSDN-GRA is presented in the presence of a percentage of malicious nodes in the network. The perspective of managing non-collaborative vehicles is briefly introduced as future work.

1 Introduction

The Vehicular Adhoc Networks (VANET) suffers from many lacks such as delays of communication because of the multi hop scenarios, frequent links failures because of the nodes mobility, and also a low security due to the several attacks and intrusions against the wireless and the heterogeneous network.
In our previous works [1][2], the SDN paradigm has been identified as a suitable approach for dealing with: the dynamic and large scale environment, the robustness of communication, the heterogeneity of the network concerning applications and the communication technologies, the routing strategies and also the security issues [3].

However, find the best way to implement the controller to suit to vehicular networks constraints is unresolved, especially because of the SDVN vulnerabilities. The security and the routing features are controlled by the SDN Controller. Whereas, the data plane applies according to the controller rules, since it is devoid of intelligence. As a result, the location choice of the controller is critical, as robust and secure communications depends on the robustness, the reliability and the availability of this controller.

This paper represents an extended version of our previous work [1] untitled "Securing Southbound Interface of HSDN-GRA Vehicular Routing Protocol using a Distributed Trust" that is published in 2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC). We described an IoV communication use case based on the SDN architecture, to root IoV's data in a secure and robust manner from a source to a destination. The proposition was a preliminary approach without simulation analysis.

In this extended paper, we firstly discuss the security vulnerabilities of vehicular networks, SDN paradigm and SDVN networks. Secondly, we propose to secure the HSDN-GRA routing protocol by using a an encrypted function and trust model [1]. In this trust model, each new vehicle broadcast its public key to its neighbors,

[*]CY Cergy Paris University, Avenue du Parc 95000 Cergy Grance, +33769265726 & lae@eisti.eu

all the keys are managed in a distributed way. Besides, the described approach aims to assign a weight ω to each vehicle. ω is deduced from the free load value as well as the trust value of each vehicle per those of their neighbors.

To deduce the trust value of a vehicle, this approach uses the error log of HSDN-GRA protocol [4], as the historical misbehaviors of a vehicle gives its trust rate. three types of unreliable behaviors that a vehicle had in a period of time are considered. The list of each vehicle's misbehavior is stored in an error log and embedded on a specific cluster head vehicle. A vehicle can be the next relay according to its weight walue.

the advantage of the proposed solution is the improvement of security in the communication process, indeed, vehicle authentication and data integrity are obtained using a signature mechanism, while the confidentiality of the data exchanged is provided by an encryption function. the main purpose of secure HSDN-GRA is to protect the routing process from malicious and unstable vehicles.

The present paper is organized as follows: Section 2 studies the security attacks against vehicular networks and SDN architecture and some security solutions for SDVN. Section 3 describes our Secure HSDN-GRA approach. The implementation details and simulation analysis is presented in Section 4. Finally, Section 5 concludes the main contributions and summarizes the perspective of an incentive scheme in order to manage the non-collaborative vehicles.

2 Background and Related Work

In the survey [2] realized by Alouache et al., the SDN architecture is described as an alternative to secure the network and to ensure the data integrity. In fact, a global overview of the system is achievable thank to the SDN controller. It collects information about the entire network traffic instead of exchanging a large amounts of information. As a result, it provides a better and reliable security mechanism than those applied to the traditional network [5].

2.1 Security Attacks in Vehicular Networks

Vehicular networks suffer from several attacks which compromise the main security requirements. [6].

In Table 1 non exhaustive list of these attacks are given.

2.2 Security Attacks in SDN Architecture

The SDN paradigm also has its own vulnerabilities, they are principally related to the characteristics, the location and the crucial role of the controller, as well as the control data flow exchanged between the control plane and the data plane [7].

Table 2 exposes a non exhaustive list of attacks related to the SDN paradigm.

2.3 Security Solutions for SDVN

Despite the vulnerabilities exposed in Table 1 and those mentioned in Table 2, and since SDN exceeds the limits of traditional mechanisms, it can be beneficial for IoV security deficiencies.

Table 1: Vehicular Network Attacks vs Security Requirements

Attacks in Vehicular Networks	Authentication	Availability	Confidentiality	Integrity	Non-repudiation
Denial of Service (DoS)		x			
Distributed DoS		x			
Jamming		x		x	
Malware attack		x	x	x	
GPS Spoofing	x	x			
Hijacking of session	x				
Position faking	x				
Illusion attack	x	x			
Bogus information attack	x	x			
GPS Spoofing	x	x			
Snooping			x		
Identity reveling			x		
Location tracking			x		
Brute force	x		x		
Eavesdropping			x		
Sink Hole attack		x	x	x	
Black Hole attack		x		x	
Masquerade attack	x			x	
Message tempering				x	x
Message suppression				x	x
Message reply	x				x
Repudiation		x			x

In fact, the SDN controller labels and isolates suspect flows and their sources, consequently, the data plane will not process the

packets coming from these flows.
Besides, the IoV network is heterogeneous and uses a set of communication technologies such: WAVE, DSRC, LTE, etc. having their own security policies, therefore, due to the SDN controller and its given global overview, these security policies can be deployed without conflict, which reinforces the security of the whole IoV system.

In addition, the centralization and the abstraction provided by the controller give the administrators the possibility to update any security policy based on observed unreliable behaviors. Table 3 describes how SDN can be a solution for security issues.

Table 2: SDN Attacks vs Security Requirements

SDN attacks in SDN Architecture	Authentication	Availability	Confidentiality	Integrity	Non-repudiation
DoS on the Data plane		x			
Distributed DoS on the Control plane		x			
Controller identity spoofing	x	x	x	x	x
Flow based forwarding attack				x	x
Flow table alteration			x	x	

Table 3: Security in IoV with SDN Architecture

Action	details
Intrusion detection	The SDN controller manages the nodes authentication and detects malicious intrusions.
Attacks identification	The SDN controller monitors the exchanged flows emanating from all the nodes to detect any setting alteration.
Self recovery mechanism	The SDN controller hosts rules to automatic recovery against any attack.

Recently, some research papers deal with both routing and security issues in vehicular networks by adopting the SDN paradigm.

The Improvised Trust based Ad-hoc On-demand Distance Vector routing (I-TAODV) is proposed by Vasudes et al. in [8], it aims to secure routing in a multi hop scenario.
Based on a metric called trust value, the authors propose an algorithm to identify the trusted vehicles, and another one to identify the malicious ones.
This protocol is built on the SDN paradigm, where the control plane monitors forwarding, reversing, trust of forwarding Vehicle, trust of reverse vehicle, path trust and network performances.

Zhang et al. elaborate The Software-Defined Trust based Ad hoc On-demand Distance Vector routing (SD-TAODV) in [9]. The process of route discovery, route maintenance, the reverse and forwarding paths are chosen by the control plane.
They also use a trust management mechanism which is represented by a bi-objective function, it tries to optimize two objectives: the node trust and the path trust calculated respectively by Trust Node Calculation Process and Path Trust Calculation Process They are used to enhance the Route Discovery Process of TAODV.

3 Securing the HSDN-GRA

The HSDN-GRA protocol presented in the research [4] exploits a log of communication errors where three types of vehicle's misbehavior are stored: link failures, random reception of Beacon messages, and non-acknowledgement of previous messages.
To estimate the trust value, the vehicle communication history is observed and tracked from the log of communication error. In this extended paper, we use the trust value of each vehicle with comparison to the trust values of all its neighbors.

The trust is described as the expectation and the belief that a vehicle has about other vehicles concerning future behaviors.
Its estimation is based on 1) experiences and evidences collected in the past either directly or indirectly, and 2) the knowledge about the vehicles nature, and/or on recommendations from trusted entities [10]-[12].

This trust value in addition to the free load value of each vehicle are aggregated to assign a weight ω to each vehicle. This weight will be used to elect the best relay at each step of the routing process.

The routing rules are commonly governed by two types of controllers in HSDN-GRA: the semi-centralized controller represented by the cluster head, and the distributed controllers represented by the cluster members.

Figure 1 shows an IoV communication scenario using the HSDN-GRA routing protocol based on a semi-centralized SDN architecture.
A vehicle tries to access to an Internet service is requested by a vehicle, but the request fails because the closest infrastructure is down. So it switches to the vehicles on the road for routing its request.

The network is divided on clusters and Cluster heads are elected. Each Cluster Head contains a part of the control plane (the error log). The second part is distributed on the rest of the cluster members.

until reaching the destination, the request transits through vehicles with a hop by hop approach. Each vehicle selects its next relay with respect to the controller policy i.e,: The one who has the best relative trust value and the best relative free load value.

Hypothesis:

- In this extended paper, we opt for the highway scenario, this choice is guided by the stability of vehicle's speed.
- At first, the free load and the trust values of vehicles are equal and optimal.
- Malicious-Vehicles $\ll Reliable - Vehicles$.

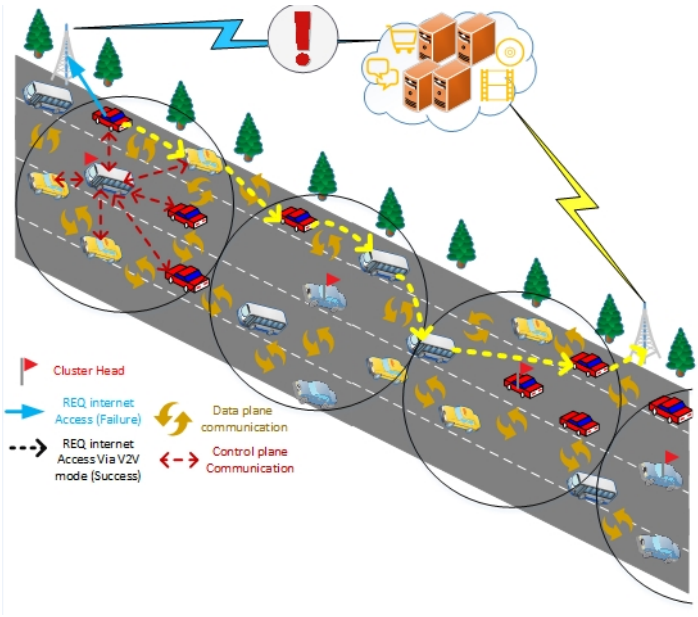


Figure 1: A Use Case of IoV Communication with a Hybrid SDN Architecture

The following subsections describe the securing HSDN-GRA steps.

3.1 Public key Distribution

- all the vehicles exchange a define Beacon to form a a one hop neighboring.
- Each vehicle sends its Public Key *Pkey* to the one hop neighbors, and receives back their *Pkey*.
- At this initialization step, because of the absence of cluster head, only the distributed controllers embedded on each vehicle are operational.
- Each vehicle has a weight ω, it exploits the load capacity l_i deduced from the received Beacons, and the trust value $Trust_i$ calculated from the log of communication errors according to Formula 1. The weight ω of each vehicle is calculated from Formula 2 by all the neighbors, and the true value will be confirmed by the majority.
- The election of the Cluster Head is realized by a distributed consensus after the initialization period of time T has passed, the one with the largest weight ω will be the Cluster Head.

$$Trust = (\frac{1}{Popularity}) \quad (1)$$

$$\omega(v)_i = (\frac{l_i}{\sum_{j=1}^{m} l_j} + \frac{Trust_i}{\sum_{j=1}^{m} Trust_j}) \quad (2)$$

where m is the number of neighbors.

-To monitor the communication errors of each neighbor, in the beginning, before electing cluster heads, logs are distributed within all the vehicles. The trust rate is estimated by exploiting all the distributed logs. So, initially each vehicle:
-Monitors its neighbors and has it own log
-Calculates the weight ω of each neighbor
-Then, each vehicle broadcasts in a map the $Pkey_i$ of each neighbor and their associated weight ω values:

$$NeighborPkeyMap = \{Pkey_i, \omega_i)\}, \text{With } i \in \{1, m\}$$

The algorithm 1 describes how the public key distribution as well as the weight of each vehicle is managed.

Algorithm 1: Public Key Distribution (PKD)

Let $Pkey_i$ be a the public key of vehicle V_i;
Let $NeighborsPkeyMap < Pkey, \omega >$ be a the map in which a vehicle V store the public key and the weight of its neighbors;
input : Vehicle V
output : $NeighborsPkeyMap$
Initialization: $NeighborsPkeyMap = \emptyset$; ω_{Vi}=*null*; $Pkey_i$=*null*;
for *each vehicle* V_i **do**
  $Broadcast(Beacon_{Vi})$;
  $Broadcast(Pkey_i)$;
end
for *each vehicle* V **do**
  $Receive(Beacon_{Vi})$;
  $Receive(Pkey_{Vi})$;
  $Trust_i = (\frac{1}{Popularity_i})$;
  $\omega(V)_i = (\frac{l_i}{\sum_{j=1}^{m} l_j} + \frac{Trust_i}{\sum_{j=1}^{m} Trust_j})$;
  $Insert(< Pkey_{Vi}, \omega(Vi) >, NeighborsPkeyMap)$;
end
$Broadcast(NeighborsPkeyMap)$;

3.2 Cluster Head Election

-Using all the received $NeighborsPkeyMap_i$, a vehicle V confirms the reliable weight value ω_i of each neighbor V_i and the association $(V_i/Pkey_i)$. i.e, a coherent contain of the $NeighborsPkeyMap$ is deduced by the majority who give the same values.
-On each vehicle a consensus occur to chose as Cluster Head the vehicle with the largest weight ω.
-The variable "IsCh" will be $True$ in the Beacon for the Cluster Head, while the rest of vehicles will keep the value $False$.
-The elected Cluster Head, representing the Semi-Centralized SDN controller, will now hosts the log of communication errors related

to it cluster.
The algorithm 2 details the Cluster Head election.

Algorithm 2: Secure Cluster Head Election (SCEA)

Let $NeighborsPkeyMap_i < Pkey, \omega >$ be a the maps received by V from each neighbor;
input : Vehicle V
output : $ClusterHead$
$Initialization$: $Max\omega = \omega$;$PkeyMax=Pkey$; $ClusterHead = null$;
for *each vehicle V* **do**
  $Receive(NeighborsPkeyMap_i)$;
  $FixTrueValueof(NeighborsPkeyMap)$ By the Majority;
  if $(NeighborsPkeyMap.\omega_i > Max\omega)$ **then**
    $Max\omega = NeighborsPkeyMap.\omega_i$;
    $PkeyMax = NeighborsPkeyMap.Pkey_i$;
  else
    if $(NeighborsPkeyMap.\omega_i == Max\omega)$ *and* $(NeighborsPkeyMap.Pkey_i < PkeyMax)$ **then**
      $Max\omega = NeighborsPkeyMap.\omega_i$;
      $PkeyMax = NeighborsPkeyMap.Pkey_i$;
    end
    $ClusterHead = V_{Max\omega}$;
    if $(ClusterHead = V)$ **then**
      $Beacon_V.isCH = True$;
      $Send(NewBeacon)$;
      $Receive(ErrorLog)$ from previous Cluster Head or from the neighbors at initialization;
      $Filter(ErrorLog)$;
    else
      $Reset(NeighborsPkeyMap_i)$;
    end
  end
end

3.3 Secure Control Plane of HSDN-GRA

In this section, the Secure HSDN-GRA is detailed.
Once the first election of the Cluster Head occurred, the vehicles signs all their Beacons. The objective is to guarantee the authentication.Figure 2 expose the Beacon message structure.

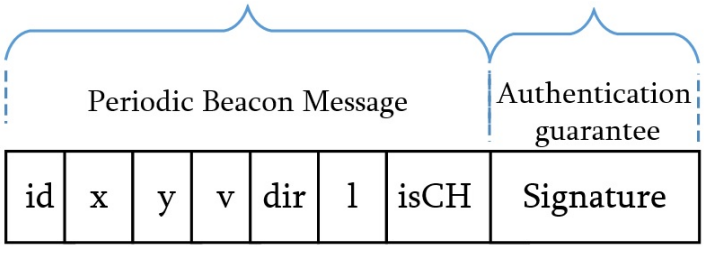


Figure 2: Beacon Message structure in Secure HSDN-GRA

The structure of the messages exchanged via the southbound interface are represented in Figure 3.

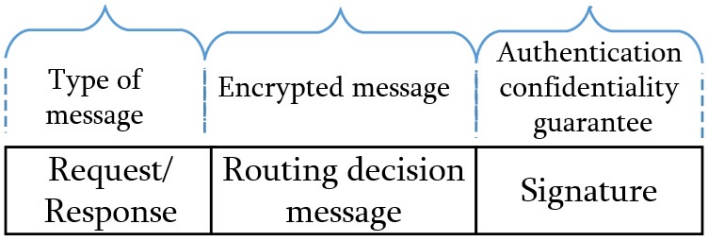


Figure 3: Structure of Southbound Messages in Secure HSDN-GRA

As the confidentiality of the control decisions, especially about the next hops, need to be preserve, we also propose to secure the southbound interface used in HSDN-GRA routing protocol.

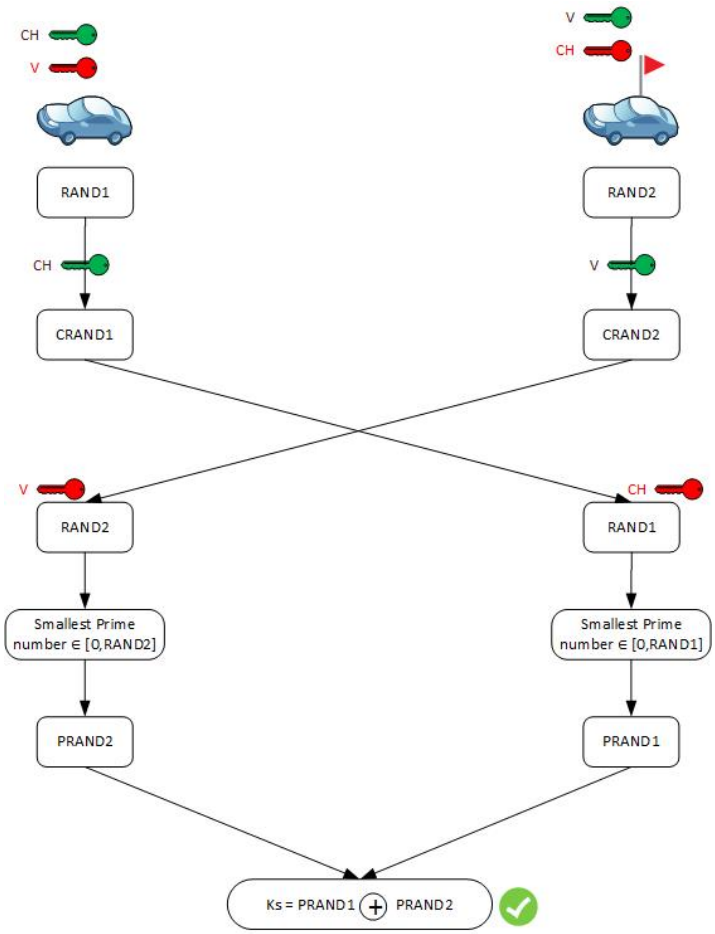


Figure 4: Secret Key creation between the Cluster Head and a vehicle V

To encrypt the data flow of the southbound interface, we are inspired from the symmetric encryption algorithm AES [13].A secret key K_s is exchanged between each vehicle and its Cluster Head. The public keys $Pkey$ already shared are exploited to deduce K_s.

A vehicle V and its Cluster Head exchange two random numbers $RAND1$ and $RAND2$ encrypted respectively with $Pkey_{CH}$ and $Pkey_V$.

Then, the Cluster Head decrypt with its private key $RAND1$

and choose the smallest prime number $PRAND1$ in the interval [0,RAND1]. The vehicle V decrypt also with its private key $RAND2$ and choose the smallest prime number $PRAND2$ in the interval [0,RAND2].

Finally, the secret key K_s is given using the formula 3. The symbol $\bigoplus$ representing the Exclusive Or logical operation.

$$Ks = (PRAND1 \bigoplus PRAND2) \quad (3)$$

Figure 4 illustrate how the secret key K_s is formed by the Cluster Head and a vehicle V in order to encrypt the southbound interface communications.

Figure 5 illustrate communications inside a cluster of Secure HSDN-GRA.

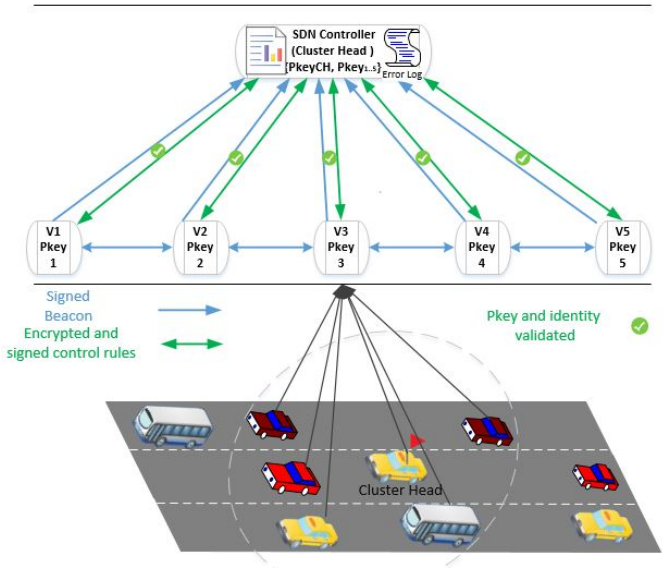


Figure 5: Communication inside a Cluster of Secure HSDN-GRA

3.4 Secure Incoming Vehicles in a Cluster

Three steps are required to integrate a new vehicle in a cluster:

- The new vehicle V_n broadcast its Beacon and its $Pkey$ to announce its self.
- All the neighbors inside the Cluster reply to V_n with the identifier and the $Pkey$ of the Cluster Head, while sending the triplet (V_n,$Pkey_n$, ω_n) to the Cluster Head.
- The Cluster Head analyzes all the received triplets (V_n,$Pkey_n$,ω_n) in order to confirm the identity of V_n. It deduces the percentage of veracity of this triplet, from the number of vehicles who affirm it. After that, a Challenge-Response mechanism [14] is exploited associate definitively V_n to $Pkey_n$.

These steps are respectively illustrated by Figure 6, Figure 7 and Figure 8.

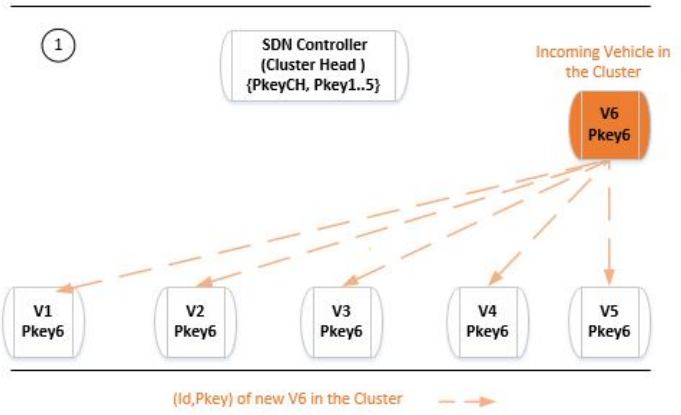


Figure 6: Incoming Vehicle in the Cluster: Phase 1

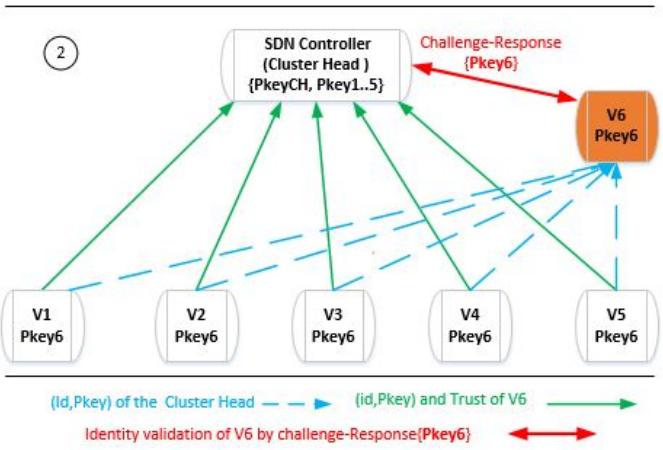


Figure 7: Incoming Vehicle in the Cluster: Phase 2

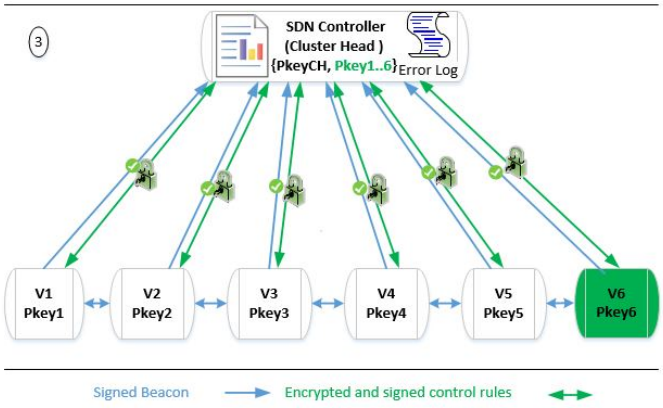


Figure 8: Incoming Vehicle in the Cluster: Phase 3

3.5 Secure Outgoing vehicle from a Cluster

This section shows the behaviour of secure HSDN-GRA in case of outgoing vehicle. Two cases are identified:

The first case deals with a cluster member which formulates a Request-To-Leave message the steps are:

- The Cluster member sends a Request-to-Leave message to the Cluster Head
- A Challenge-Response mechanism [14] is launched to confirm that the request is formulated by the pretending cluster member.
- After the authentication, an ACK message is sent by the Cluster Head to the leaving vehicle as an acknowledgement of its request. Meanwhile, the Cluster Head also broadcasts a $Revoke(V, Pkey)$ message inside the cluster in order to delete the keys of the outgoing vehicle.
- ACK messages confirming the revocation of this vehicle are sent to the Cluster Head by all the cluster members.

Algorithm 3: Secure Outgoing Vehicle From a Cluster (SOC)

Let τ be the Beacon interval and ToF be the maximum delay of receiving Beacons;
input : Vehicle V
output : $Cluster$
$Initialization$: $ToF = 5\tau$;
for *each Cluster* **do**
  if *(V wants to leave a Cluster)* **then**
    $Send(Request - To - Leave)$ to the Cluster Head);
    if *($Challenge - Response(ClusterHead, V)$ is verified by the Cluster Head)* **then**
      $Send(ACK)$ by the Cluster Head to V;
      $Broadcast(Revoke(V, Pkey))$ to the cluster members;
    end
  else
    if $(Receive(Beacon_V) \notin [\tau,5\tau])$ **then**
      $Broadcast(Revoke(V, Pkey))$ to the cluster members;
    else
      $Reset(ToF)$;
    end
  end
end

The second case concerns the revocation of the vehicles that don't respect the periodicity of Beacons, as they are considered as unreliable.
In this paper, at every time-out τ, vehicles are supposed to send and receive back Beacons. Besides, a Timer-of-Refresh ToF, initialized to 5τ, is assigned to each Cluster member sending a Beacon. The Cluster Head launches and decrements ToF when it receives a new Beacon message from a Cluster member.

If the Cluster Head receives t Beacon in the time interval $[\tau,5\tau]$, the ToF value is refreshed and reset to its initial value. Otherwise, the Cluster member is considered as unreliable and need to be excluded from the cluster according to the following steps:

- If the ToF has passed, the Cluster Head broadcasts a $Revoke(V, Pkey)$ message within its cluster.
- ACK messages confirming the revocation of this vehicle are sent by all the Cluster members.

The two cases are combined in the algorithm 3.

Figure 9 shows a Cluster member outgoing scenario.

It encompasses the two cases. The black circle details the Request-to-Leave scenario, while the rest explains the revocation procedure.

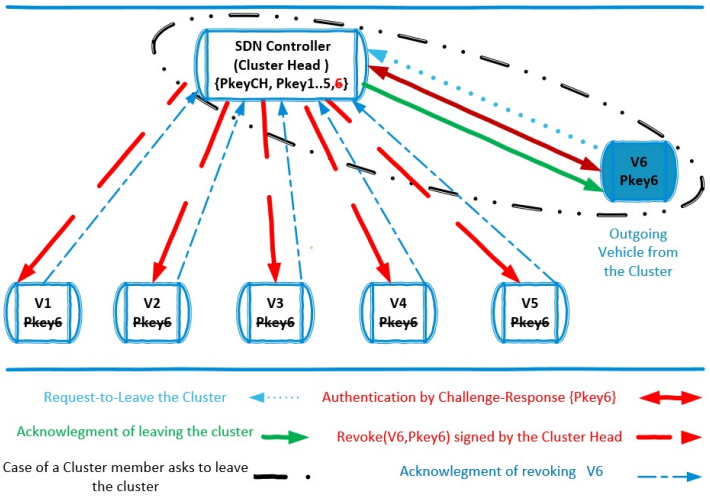


Figure 9: Outgoing Cluster Member Management

4 Simulation Analysis

In this simulation, we simulate a scenario of the Secure HSDN-GRA routing packets in IoV.
We develop a simulation with traffic condition similar to the situation in Figure 1.

Table 4: NS2 Simulation Parameters

Parameters	Specification
Simulation time	300s
Simulation area	1000m X 1000m
Number of nodes	[20-300]
Speed	[50-140] km/h
Propagation model	Two Ray Ground
Medium capacity	6 Mbps
Transmission range	310 m
Transport layer	UDP

4.1 NS2 Implementation

The simulation is done under a Ubuntu 10.04 Linux machine where we have installed the new version of network simulator NS2.34[15] and VanetMobiSim[16].
The simulation parameters are summarized in Table 4.

Some simulation functions in C++ are summarized in listing 1 (cf. Appendix 5).

4.2 Experimental Results

The goal of this evaluation consist of the study of attack resilience and routing performances of the secure HSDN-GRA protocol against the insecure HSDN-GRA protocol.

Percentage rate of data delivery and average end to end delay are chosen as indicators of routing performances.

The comparative study is done for different combinations of metrics like: the presence of malicious nodes, the clusters density and the vehicles speed.

the routing decisions of HSDN-GRA protocol are guided by the one-hop neighbors, besides, the secure HSDN-GRA protocol aims to detect and avoid malicious nodes. So, the routing performances will be positively affected by the improvement of routing behavior in presence of malicious nodes

4.2.1 Effect of Malicious Nodes

In this section, we observe in Figure 10 that the Secure HSDN-GRA presents better packet delivery ratio than the baseline HSDN-GRA once an increased number of malicious nodes are introduced in the simulation.

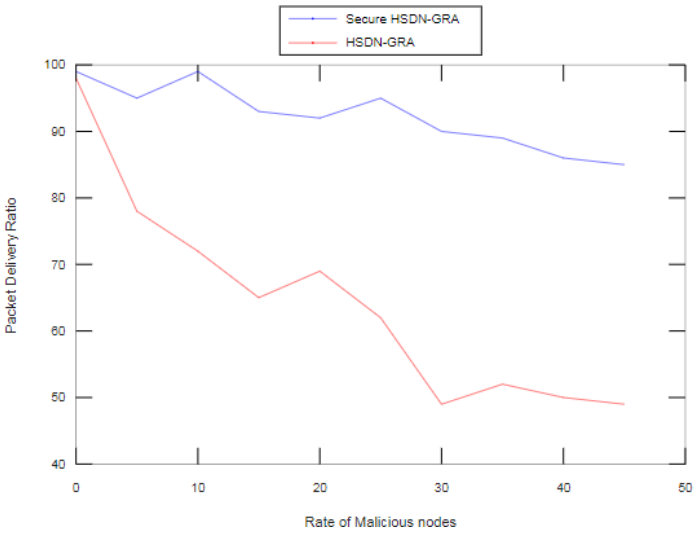


Figure 10: Packet Delivery Ratio vs. Number of Malicious Nodes

Indeed, the Secure HSDN-GRA elects as a relay the neighbor with the best trust value as well as the maximum free load with comparison to the whole neighbors, furthermore, the exchanged packets are signed and encrypted with a secret key Ks which makes the alteration and the destruction of packets more difficult.

Besides, the Secure HSDN-GRA also outperforms the baseline HSDN-GRA of delay in the presence of an increased number of malicious nodes in the simulation as shown in Figure 11.

In fact, the Secure HSDN-GRA takes into account the trust value of each relay with comparison to the whole neighbors as well as the corresponding average free load before choose it. It makes the HSDN-GRA more resilient.

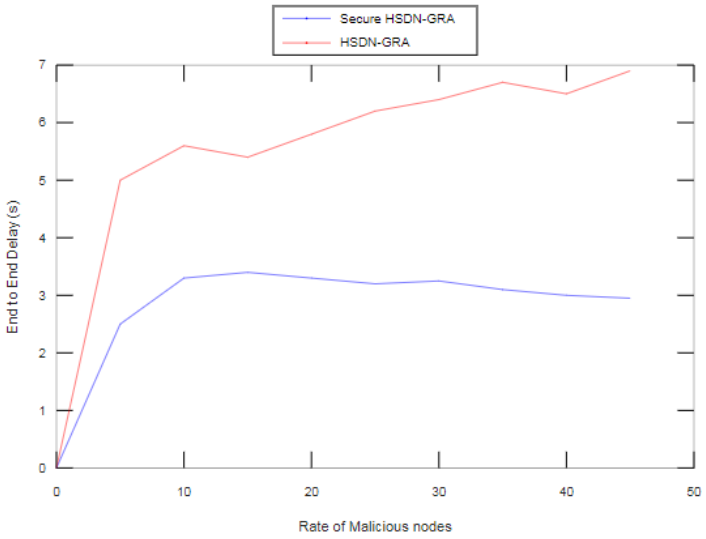


Figure 11: Average End to End Delay vs. Number of Malicious Nodes

4.2.2 Effect of Cluster nodes Density

Figure 12 and Figure 13 show respectively the impact of clusters density on the behaviour of both Secure HSDN-GRA and baseline HSDN-GRA in term of packet delivery ratio and delay especially in the presence of 20% of malicious nodes.

Firstly, for both Secure HSDN-GRA and baseline HSDN-GRA the increase density inside clusters offer better packet delivery ratio because more relays can be exploited. However, the packet delivery ratio is more satisfying when the relays are chosen according to the trust analysis estimated by the network.

Then, we observe in Figure 12 that Secure HSDN-GRA deliver an average of 30.28% of packets more that baseline HSDN-GRA in the presence of 20% of malicious nodes.

Besides, we can deduce that the authentication and the encrypted communications decrease the number of packets alteration, packet lost and malicious intrusion.

Secondly, we can observe that Secure HSDN-GRA reduce the delay comparing to baseline HSDN-GRA because the relays are chosen according to their trust values in order to guarantee the local acknowledgement of each packet sent and reduce the error recovery, consequently the delay is reduced.

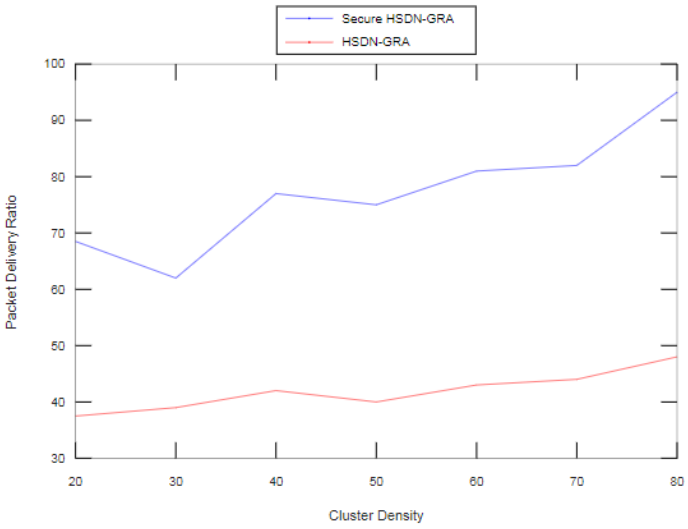


Figure 12: Packet Delivery Ratio vs. Clusters Density with 20% of malicious nodes

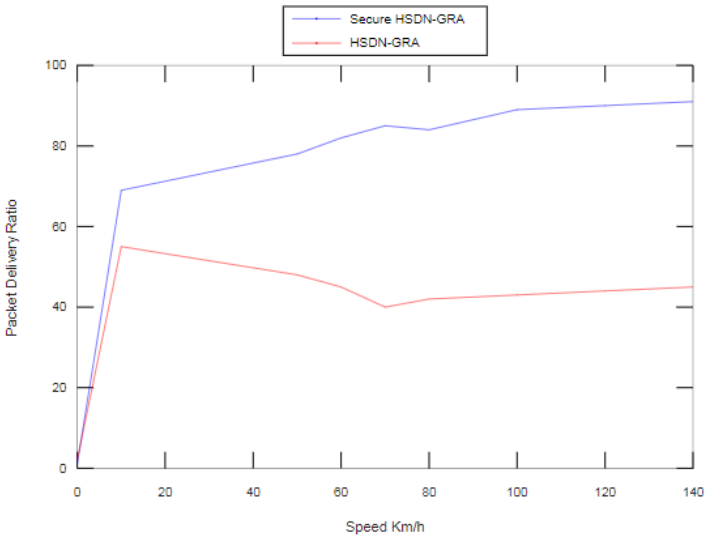


Figure 14: Packet Delivery Ratio vs. Speed with 20% of malicious nodes

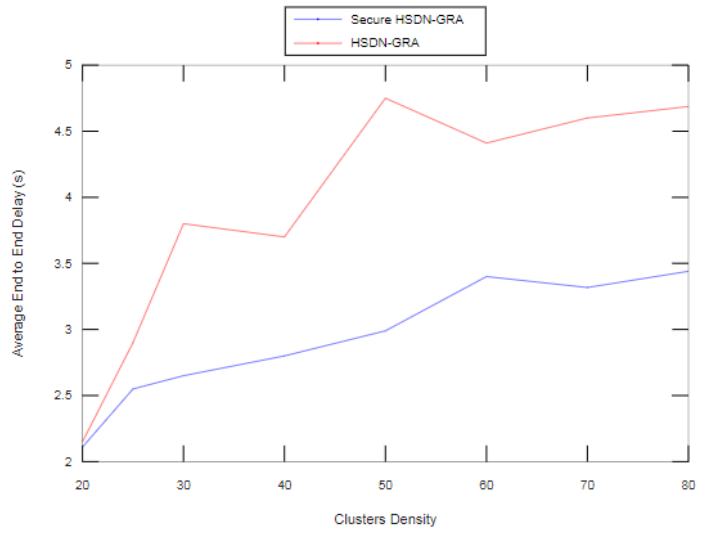


Figure 13: Average End to End Delay vs. Clusters Density with 20% of malicious nodes

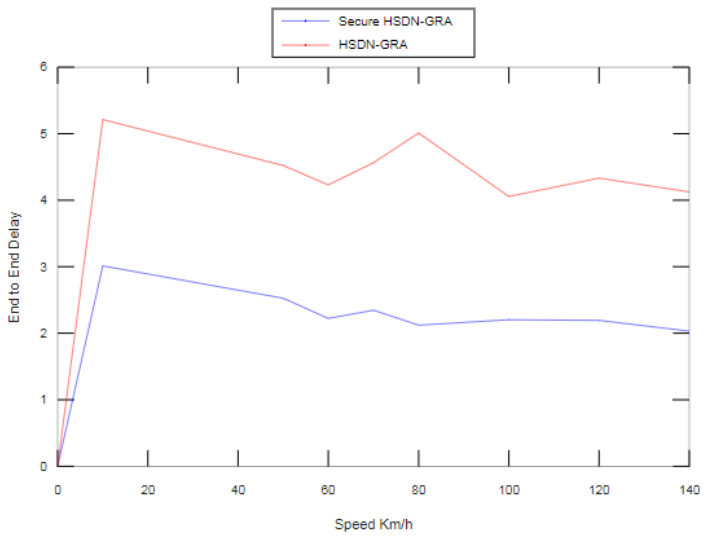


Figure 15: Average End to End Delay vs. Speed with 20% of malicious nodes

4.2.3 Effect of Vehicles Speed

Figure 14 and Figure 15 show the impact of vehicles speed against the Secure HSDN-GRA and baseline HSDN-GRA in the presence of 20% of malicious nodes.

The results in Figure 14 show that the secure HSDN-GRA outperform the baseline HSDN-GRA in term of packet delivery ratio, besides, we can observe that the increase speed stabilizes the packet delivery ratio for the both protocols.

For the delay performances, we can observe in Figure 15 that Secure HSDN-GRA reduces the delay when the speed, furthermore, it surpasses the baseline HSDN-GRA, it is due to the stability of the clusters on the highways and the packets destination are reached faster, since the relays are chosen according to the direction of the destination.

5 Conclusion

The result of applying the SDN architecture on vehicular network consist of making the routing strategies optimal, especially for with load balancing, security policies and also for the network heterogeneity management.

Indeed, it aims to select the most suitable channels and frequencies for data transmission at a specific time according to the context and the requirements.

Besides, it implements various security policies and exploits them adaptively according to the requirements. Finally,the nodes are balanced thanks to the global view given by the control plane.

In this paper, a secure HSDN-GRA for more robustness is exposed, a trust model is built between nodes, and the communications

between the nodes as well as those of the southbound interface are secured with an encryption function.
Even if we have a performing and securing routing protocol, in a totally distributed system it is difficult to guarantee the cooperation of all the entities without any rewards in turn, the selfish nodes constitute an obstacle of communication in a totally distributed network, so as a future perspective, a solution based on blockchain paradigm will be integrated in secure HSDN-GRA in order to tackle selfish nodes and stimulate their cooperation.

Acknowledgment This project was supported by Initiative d'Excellence Paris//Seine.

References

[1] Lylia Alouache et al. "Securing Southbound Interface of HSDN-GRA Vehicular Routing Protocol using a Distributed Trust". In: *Fourth International Conference on Fog and Mobile Edge Computing, FMEC 2019, Rome, Italy, June 10-13, 2019*. 2019, pp. 90–97.

[2] Lylia Alouache et al. "Survey on IoV routing protocols: Security and network architecture". In: *International Journal of Communication Systems* 32.2 (2019), pp. 38–49.

[3] Antonio Di Maio et al. "Enabling SDN in VANETs: What is the Impact on Security?" In: *Sensors* 16.12 (2016).

[4] L Alouache et al. "Toward a hybrid SDN architecture for V2V communication in IoV environment". In: 2018 Fifth International Conference on Software Defined Systems (SDS). Barcelona, Spain, 2018, pp. 93–99.

[5] P. Baskett et al. "SDNAN: Software-Defined Networking in Ad-Hoc Networks of Smartphones". In: *2013 IEEE 10th Consumer Communications and Networking Conference (CCNC)*. Las Vegas, NV, USA, 2013, pp. 861–862.

[6] M. Azees, P. Vijayakumar, and L. Jegatha Deborah. "Comprehensive survey on security services in vehicular ad-hoc networks". In: *IET Intelligent Transport Systems* 10.6 (2016), pp. 379–388.

[7] M. Liyanage et al. "Opportunities and Challenges of Software-Defined Mobile Networks in Network Security". In: *IEEE Security Privacy* 14.4 (2016), pp. 34–44.

[8] H. Vasudev and D. Das. "A trust based secure communication for software defined VANETs". In: *2018 International Conference on Information Networking (ICOIN)*. 2018, pp. 316–321.

[9] D Zhang et al. "Software-defined Vehicular Ad Hoc Networks with Trust Management". In: *the 6th ACM Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications*. DIVANet '16. Malta, Malta: ACM, 2016, pp. 41–49. ISBN: 978-1-4503-4506-4.

[10] Seyed Soleymani et al. "Trust Management in Vehicular Ad-Hoc Network: a Systematic Review". In: 2015 (2015).

[11] N. Bimeyer et al. "Assessment of node trustworthiness in VANETs using data plausibility checks with particle filters". In: *2012 IEEE Vehicular Networking Conference (VNC)*. 2012, pp. 78–85.

[12] Junhai Luo, Xue Liu, and Mingyu Fan. "A Trust model based on fuzzy recommendation for Mobile Ad-hoc Networks". In: *Computer Networks* 53.14 (2009), pp. 2396 –2407. ISSN: 1389-1286.

[13] Jim Schaad. *Use of the Advanced Encryption Standard (AES) Encryption Algorithm in Cryptographic Message Syntax (CMS)*. RFC 3565. 2003. URL: `https://rfc-editor.org/rfc/rfc3565.txt`.

[14] Margaret Rouse. *Challenge - Response Authentication Definition: What does Challenge - Response Authentication mean?* `https://searchsecurity.techtarget.com/definition/challenge-response-system`. Accessed: 2018-05-30.

[15] Teerawat Issariyakul and Ekram Hossain. *Introduction to Network Simulator NS2*. 1st ed. Springer Publishing Company, Incorporated, 2008. ISBN: 0387717595, 9780387717593.

[16] Jrme Hrri et al. "Vehicular mobility simulation with VanetMobiSim". In: *SIMULATION* 87.4 (2011), pp. 275–300.

Appendix

Table 5 presents the abbreviations used throughout the paper.

Table 5: Abbreviations used in the paper

ACK	Acknowledgment
AES	Advanced Encryption Standard
DoS	Denial-of-Service
DSCR	Dedicated Short-Range Communication
FMEC	International Conference on Fog and Mobile Edge Computing
GPS	Global Positioning System
HSDN-GRA	Hybrid SDN-based Geographic Routing Protocol with Multi-agent Approach
IoV	Internet of Vehicle
I-TAODV	Improvised Trust based Ad-hoc On-demand Distance Vector routing
LTE	Long Term Evolution
SDN	Software-Defined Networking
SD-TAODV	Software-Defined Trust based Ad hoc On-demand Distance Vector routing
SDVN	Software-Defined Vehicular Networking
TAODV	Trust based secure routing in AODV routing protocol
ToF	Timer-of-Refresh
UDP	User Datagram Protocol
V2V	Vehicle-to-Vehicle
VANET	Vehicular Ad Hoc Network
WAVE	Wireless Access in Vehicular Environments

Listing 1: Some NS2 Simulation Functions

```
static inline double jitter (double max, int be-random-)
static class ANGLEclusterHeaderClass : public PacketHeaderClass
static class ANGLEclusterClass
ANGLEclusterClass() :
TclClass("Agent/ANGLEcluster"):
public TclClass ("PacketHeader/ANGLEcluster",sizeof(hdr-all-
    anglecluster))
void ANGLEclusterHelloTimer::expire(Event *e)
UpdPosTimer::expire(Event *e)
StateTimer::expire(Event *e)
PERTimer::expire(Event *e)
EndElecTimer::expire(Event *e)
ANGLEclusterAgent::ANGLEclusterAgent() : Agent(PT-ANGLECLUSTER)
int ANGLEclusterAgent::command(int argc, const char*const* argv
    )
void ANGLEclusterAgent::SendHello()
void ANGLEclusterAgent::SendBeacon(u int8 t type)
void ANGLEclusterAgent::SendMSG()
void ANGLEclusterAgent::SendRoutingMSG(nsaddrt s,nsaddrt d, int
     i)
void ANGLEclusterAgent::sendACK(nsaddrt node, float seqnumber)
void ANGLEclusterAgent::sendPkey()
void ANGLEclusterAgent::sendRequestToLeave(nsaddrt node)
void ANGLEclusterAgent::sendRevoke(nsaddrt node)
void ANGLEclusterAgent::receive(Packet* p, Handler*)
void ANGLEclusterAgent::RecvHello(Packet *p)
void ANGLEclusterAgent::RecvBeacon(Packet *p)
void ANGLEclusterAgent::RecvCHBeacon(Packet *p)
void ANGLEclusterAgent::RecvRoutingMsg(Packet *p)
void ANGLEclusterAgent::ReceiveACK(Packet *p)
void ANGLEclusterAgent::ReceiveRevoke(Packet *p)
void ANGLEclusterAgent::timeout()
tosend ANGLEclusterAgent::caluclWeight()
int ANGLEclusterAgent::search(nsaddrt inode)
void ANGLEclusterAgent::getSpeed()
void ANGLEclusterAgent::getmovspeed(double * sp)
void ANGLEclusterAgent::getLoc()
void ANGLEclusterAgent::GetLocation()
void ANGLEclusterAgent::updcoord()
void ANGLEclusterAgent::signalPre()
void ANGLEclusterAgent::signalRouting()
void ANGLEclusterAgent::showState()
void ANGLEclusterAgent::PurgeTable(u int8t num)
void ANGLEclusterAgent::secretKey()
void ANGLEclusterAgent::signature(Packet* p, Handler*)
void ANGLEclusterAgent::crypt(Packet* p, Handler*)
int ANGLEclusterAgent::numofNeighbors()
int ANGLEclusterAgent::existmember()
void ANGLEclusterAgent::Popularity(nsaddrt node,int cas)
double ANGLEclusterAgent::Frequency(nsaddrt node,int cas)
int ANGLEclusterAgent::getPopularity(nsaddrt node)
double ANGLEclusterAgent::getFrequency(nsaddrt node)
nsaddrt ANGLEclusterAgent::unReliableNode()
float ANGLEclusterAgent::getTrust(nsaddrt node)
void ANGLEclusterAgent::updatePositions()
void ANGLEclusterAgent::endelection()
void ANGLEclusterAgent::VehicleState()
void ANGLEclusterAgent::RecvPresMSG(Packet *p)
void ANGLEclusterAgent::BestNode(nsaddrt node, double pd)
nsaddrt ANGLEclusterAgent::NextRelay()
void ANGLEclusterAgent::Revoke(nsaddrt node)
```

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 614-620 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Determinants of Students' Actual use of the Learning Management System (LMS): An Empirical Analysis of a Research Model

Charles Buabeng-Andoh[*,1], Charles Baah[2]

[1]*University of Education, Winneba, Department of ICT Education, Winneba, Ghana*

[2]*Pentecost University College, Department of IT, Ghana*

ARTICLE INFO

Article history:
Received: 02 January, 2020
Accepted: 27 March, 2020
Online: 14 April, 2020

Keywords:
Learning Management System
Unified Theory of Acceptance and Use
Structural Equation Model

ABSTRACT

This study built and tested a research model to determine the determinants that impact students' actual use of the learning management system. A survey questionnaire was used to gather data from 148 university students who used LMS in their course of study. The structural equation model was used to analyze quantitative data. The study revealed that (1) performance expectancy, effort expectancy and institutional support positively impacted students' actual use of LMS, (2) social influence and infrastructure support did not positively impact students' actual use, (3) gender had a significantly moderated effect on the correlation between institutional support and actual usage of technology. This study added to existing studies on the use of UTAUT in explaining students' actual use of technology in developing nations. Implications for practice, drawbacks and future directions are discussed.

1. Introduction

Information technology has caused a series of changes in higher institutions by transforming the previous idea of learning [1]. The LMS is among the evolving technological tools that supports e-learning. LMS is a web-based system that facilitates web-based interaction between teachers and students [2], offers educators and trainers kits to develop and present study materials, check student involvement and assess learner performance, provides students with the means of networking with peers via chat, group conversation and video conferencing [3]. The most common LMS used in schools include Modular Object-Oriented Dynamic Learning Environment (MOODLE), Blackboard, WebCT and Desire2Learn [4]-[5]. Several higher learning institutions have spent millions of dollars in ICT facilities to aid the quality of the instruction process [6]-[7]. However, this capital spending will be useless if students do not accept the system [8]. A meta-analysis of LMS acceptance and use in sub-Saharan Africa (SSA) from 2007-2017 by [9], revealed a lack of research on LMS use in SSA. [9] proposed that further study should be conducted on the determinants of LMS use and acceptance in SSA by adopting the Unified Theory of Acceptance and Use of Technology (UTAUT) model [9]. Additionally, [9] suggested that a more rigorous numerical analysis such as the structural equation model (SEM) be used to determine the factors that influence LMS usage. Therefore, it is critical to examine the determinants of LMS use and acceptance in SSA, particularly in Ghana.

This present study was conducted at Pentecost University College (PUC). The University College is found in Accra, Ghana, and has 2,500 students and 100 lecturers. In 2012, the University College implemented MOODLE LMS to supplement its standard face-to-face teaching. The MOODLE offers instructors with attributes such as course and user-management tools, discussion forums, quizzes and announcements to support their coaching experience. It also offers support in the management of the e-learning process by storing learners' information organized, making course materials accessible to learners, tracking learners' progress and generating reports, supporting collaboration with learners and providing testing and evaluation tools. It also allows learners to submit assignments. The MOODLE is used to support the teaching and learning process in any subject area. The system supports contents such as text, image, audio, and video to facilitate teaching and learning. Instructors register with the system to use it. After registration, the system validates the instructor's email address, approves it and activates it. Though

[*]Corresponding Author; Charles Buabeng-Andoh, E-mail cbandoh@hotmail.com

MOODLE LMS was introduced to deliver online courses to students in the institution, most of the lecturers and students have not adopted the system. Hence, for the successful use of the LMS, it is important to understand the determinants that impact the real usage of the LMS. Studies have revealed that the real implementation of LMS has not been fully studied in SSA nations [9], and even, the scanty studies on LMS have focused on technical issues ignoring inquiries on the determinants that impact users' real application of LMS [10]. These identified gaps have motivated the researchers to research into the determinants that affect the learners' actual application of LMS.

The objectives of this work are in two folds: (1) to explore the determinants that are significant for the use of e-learning system, in precise, the LMS in a Ghanaian university and (2) to build a research model integrating the factors that are directly essential for e-learning usage as well as to determine its efficacy in explaining students' actual use of LMS.

This manuscript is structured as follows: Initially, the theoretical model is described. Second, the new research model is discussed. Third, hypotheses are developed. Fourth, the research method (population, data collection, measures, and data analysis) is described. Fifth, the results of the study are explained. sixth, the conclusion is drawn. Finally, drawbacks and future studies are discussed.

2. Theoretical Framework

2.1. The UTAUT

The UTAUT model, [11], is prevalent in studies relating to online learning. As shown in figure 1, the UTAUT model consists of four key constructs that influence the users' use of the technology. These constructs are performance expectancy (PE), effort expectancy (EE), social influence (SI) and facilitating conditions (FC). Additionally, the model consists of four moderating constructs. These are gender, age, experience and voluntariness of use [11]. Regarding the rate of application, UTAUT lags behind Technology Accepted Model (TAM) indicating that for the past decade, UTAUT is yet unpopular in reports in SSA [9]. Furthermore, a current study conducted on the use of technology, revealed that findings from studies using UTAUT might differ in diverse cultural settings [12], hence, it is essential to investigate UTAUT in diverse cultural backgrounds to augment the current literature on UTAUT. Therefore, this study used the UTAUT model to explain the determinants of students' actual use of LMS in universities in SSA.

3. The New Research Model

The main variables of the new research model were PE, EE, SI, IC and IS. The moderating variables were gender and age. In the new research model, the FC construct was broken into (1) IC and (2) IS. This is to determine the exact variable of facilitating condition construct which impacts students' actual use of LMS. In the new model, 'actual use' was used as an endogenous variable because of students' real application of LMS. Voluntariness and experience of use variables were excluded as explained by [13]. The introduction of IC, IS and actual use variables differentiate the new research model from the original UTAUT (see Fig 1). Therefore, IC, IS, PE, EE and SI explain the level of the actual use of students' LMS. The research model for this study is shown in Figure 2 below.

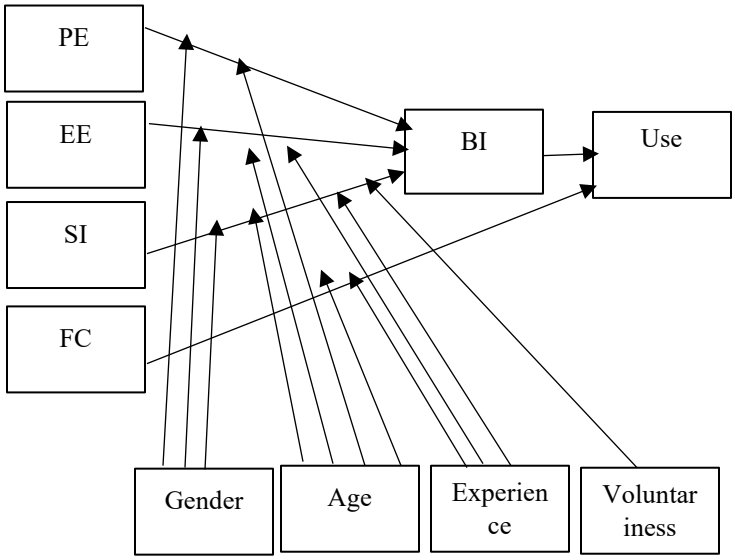


Figure 1: UTAUT model [11]

Note: BI stands for behavioral intention

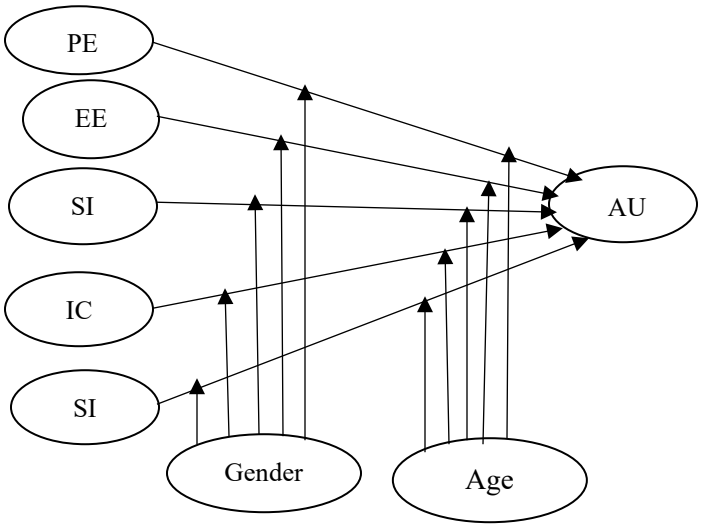


Figure 2: New study model

- PE is the belief that the performance of the user will increase with LMS use. [11].
- EE: is the trust that LMS will be easy to use [11].
- SI is a trust held by important people in the society that LMS in question should be used [11].
- IC is the belief that the technological facilities that support students to use the LMS exist.
- IS is the belief that the technological assistance the institutions offer exist.

- Actual use (AU): students' real application of the LMS.
- Gender: The gender of the student.
- Age: The age of the student.

4. Hypothesis Development

4.1. Performance Expectancy

PE is the belief that the performance of the user will increase with LMS usage [11]. According to [13] PE significantly related to technology usage. Similarly, [14] found that PE substantially related to the usage of a web-based tutoring system. Another study by [15] revealed that PE positively related to technology use. In related studies, it is shown that PE significantly impacts behavioral intention more strongly for men [11], [16]-[18]. Similarly, [11], [17], [19] have found that the older person drifts to use technology minimizes. Considering these pieces of evidence, the following hypotheses were deduced:

H1a: PE would have a significant effect on the AU of LMS.

H1b: The PE significant effect on AU of LMS would be higher when the student is male.

H1c: The PE significant effect on AU of LMS would be higher when the student is younger.

4.2. Effort Expectancy

EE is the trust that the system will be easy to use [11]. According to [20], EE significantly influences the use of technology. Also, [21] revealed that EE significantly related to technology usage. In a related study, [13] found out that EE had a significant effect on students' use of LMS. In another study, [11], [17] have posited that age and gender play a moderating role in EE and users' behavioral intention. Based on the literature above, the under-listed hypotheses were developed:

H2a: EE would have a significant effect on the AU of LMS.

H2b: The EE significant effect on AU of LMS would be higher when the student is female.

H2c: The EE significant effect on AU of LMS would be higher when the student is young.

4.3. Social Influence

SI is the extent one deems it important to use technology because a colleague or superior has asked him to use the technology [11]. Researchers, including [22] postulates that LMS usage ought not only to address the high-tech issues but also to consider the social environments where the technology is used. Also, [22] found that there is scant knowledge about the social issues regarding technology application in developing nations. In a similar work, [23] stated that SI positively influenced technology use. Moreover, a study by [24] on learners' LMS implementation revealed that SI significantly impacted AU. In a comparable work, [25] confirmed that a significant correlation exists between SI and LMS usage intention. Furthermore, [8], [17] have reported that females are more inclined to be influenced by their colleagues to use technology than males. Additionally, [16] found that age moderates significantly on the association that exist between SI and behavioral intention to use technology.

From the evidence outlined above, the hypotheses below were deduced:

H3a: SI would have a significant effect on the AU of LMS.

H3b: The SI significant effect on AU of LMS would be higher when the student is female.

H3c: The SI significant effect on AU of LMS would be higher when the student is older.

4.4. Facilitating Condition (FC)

FC is the level at which one believes that institutional and technical support are present to implement the system [11]. Training, support, institutional policy, ICT infrastructure, and leadership are the key elements considered to affect FC [26]. Many scholars have studied the impact of FC on LMS usage. [27] for instance, found out that technical support and lack of suitable equipment, and other resource availability issues significantly impacted LMS usage. [28] on the other hand, postulated that making computers and computer networks available to students will not necessarily influence usage but also administrative and institutional support should be in place to facilitate the system use. Numerous academicians also have stated the importance of support and training as core FC in e-learning use [29]-[30]. Besides, [9] has confirmed that FC significantly impacts students' usage of e-learning resources. [16] has also concluded in his study that age and gender significantly moderate FC and users' intention to use technology. In this study, FC was divided into IC and IS. From the above evidence, the following hypotheses were formulated:

H4a: IC would have a significant effect on the AU of LMS.
H4b: The IC significant effect on AU of LMS would be higher when the student is young.
H4c: The IC significant effect on AU of LMS would be higher when the student is female.
H5a: IS would have a significant effect on AU of LMS
H5b: The IS significant effect on AU of LMS would be higher when the student is female.
H5c: The IS significant effect on AU of LMS would be higher when the student is young.

5. Method

In the present study, a case study approach was adopted based on the recommendation that it is a suitable strategy for studying

technological intervention in education [31]. Moreover, [32] states that case studies enable researchers to learn more from a particular situation.

5.1. Population and Data Collection Procedure

A total of 160 students took part in this study. The students were enrolled in undergraduate programs. A simple random sampling procedure was applied to choose the respondents from all the school year levels (Year 1 to Year 4) who were studying Business and Science courses. Before administering the questionnaires to the participants, they were told that their involvement was optional, and their anonymity was guaranteed. Additionally, the respondents were informed of the purpose of the study before administering the online questionnaire to them. The online questionnaires were completed after the respondents had taken a semester course. To enhance the rate of response, students were regularly reminded to complete the online questionnaires. Overall, 148 questionnaires were received. Of the 148 partakers, 69% were males and 31% were females, 23% were in the age range of 17 and 22 whereas 77% were more than 23 years, 40% of students were in Year 3 while 26% were in Year 2 (see Table 1). According to [33], a sample size between 100 and 150 cases is ideal for the Partial least square structural equation model (PLS-SEM). Consequently, a sample size of 148 students in this study was suitable for analysis using PLS-SEM.

5.2. Measures

The variables used in this study were taken from the earlier research works. The observed variables for EE, PE, SI, IS and IC were adapted from [11]. The PE construct contained four items; the EE composed of four variables; the SI had two variables. IC and IS, each comprised of five variables. A 5-point Likert scale was adopted and started from 1 representing "strongly disagree" to 5 representing "strongly agree". On average, each partaker used at least 30 minutes to complete the questionnaire.

5.3. Data Analysis

PLS-SEM was used to analyze the quantitative data gathered from the respondents. PLS-SEM has been used in a broad array of disciplines such as Information Systems [34]. However, literature is lacking in the application of PLS-SEM in studies relating to LMS usage, particularly, in SSA [9]. Hence the use of PLS-SEM for the data analysis. The two-step approach of [35] was used for the PLS-SEM analysis. The first step measures the measurement model which comprises convergent and discriminant validities. The second step calculates the structural model.

6. Results

6.1. Measurement Model Assessment

6.1.1 Convergent Validity

The convergent validity was calculated on three stages: item reliability, composite reliability (CR) and average variance extracted (AVE) [36]. The item reliability is measured by the factor loading of an item of the construct in the model. [33] posits that a factor loading of a construct is judged reliable if its value is equal to or higher than 0.5. Table 2 shows that the factor loadings of each construct's items ranged between .56 and .93 confirming [33]'s assertion. Again, [33] postulates that, for a construct to be reliable, both its CA (α), as well as CR, should be greater than 0.7. From Table 2, CA as well as the CR of all the constructs exceeded 0.7, confirming the assertion of [33]. A further test of convergent validity, the AVE was conducted. Table 2 revealed that the AVEs of the constructs surpass 0.5 inferring that the convergent validity is acceptable [36].

Table 1: Demographics of the Respondents

Item	Type	Count	Proportion
Gender	Male	102	69%
	Female	46	31%
Total		148	100%
Age	17-22 (Young)	34	23%
	>=23 (Old)	114	77%
Total		148	100%
School year	First Year	34	23%
	Second Year	26	18%
	Third Year	59	40%
	Fourth Year	29	20%
Total		148	100%

Table 2: Convergent Validity

	AU	EE	IC	IS	PE	SI	CA	CR	AVE
AU1	**.90**	.47	.21	.35	.44	.22	.87	.92	.79
AU2	**.86**	.34	.18	.35	.38	.23			
AU3	**.91**	.47	.19	.33	.39	.20			
EE1	.35	**.76**	.41	.29	.46	.32	.79	.86	.62
EE2	.31	**.70**	.31	.15	.45	.30			
EE3	.41	**.83**	.34	.35	.55	.30			
EE4	.43	**.84**	.41	.34	.48	.35			
IC1	.16	.43	**.77**	.45	.29	.37	.72	.84	.64
IC2	.14	.35	**.78**	.32	.28	.31			
IC3	.21	.35	**.84**	.39	.28	.28			
IS1	.19	.21	32	**.56**	.21	.33	.75	.83	.50

IS2	.30	.31	.37	**.78**	.20	.24			
IS3	.31	.30	.36	**.71**	.22	.33			
IS4	.24	.19	.31	**.72**	.16	.34			
IS5	.30	.27	.35	**.76**	.17	.34			
PE1	.39	.54	.28	.20	**.78**	.31	.78	.86	.60
PE2	.26	.41	.25	.12	**.66**	.18			
PE3	.41	.50	.33	.25	**.87**	.50			
PE4	.31	.47	.22	.24	**.77**	.46			
SI1	.24	.39	.39	.43	.46	**.93**	.79	.90	.82
SI2	.19	.34	.31	.36	.41	**.88**			

6.1.2 Discriminant Validity

Discriminant validity is measured by the following criteria: (1) Each construct's loadings should exceed the cross-loadings of all the constructs [37]-[38], (2) the square roots of AVE of each construct should surpass the biggest squared correlations between any other construct [36]. Table 3 depicts the square roots of AVE (in bold) of each construct.

Table 3: Discriminant Validity

	AU	EE	IC	IS	PE	SI
AU	***0.89***					
EE	0.48	***0.78***				
IC	0.22	0.47	***0.80***			
IS	0.39	0.37	0.48	***0.71***		
PE	0.45	0.62	0.35	0.27	***0.77***	
SI	0.24	0.4	0.39	0.44	0.48	***0.91***

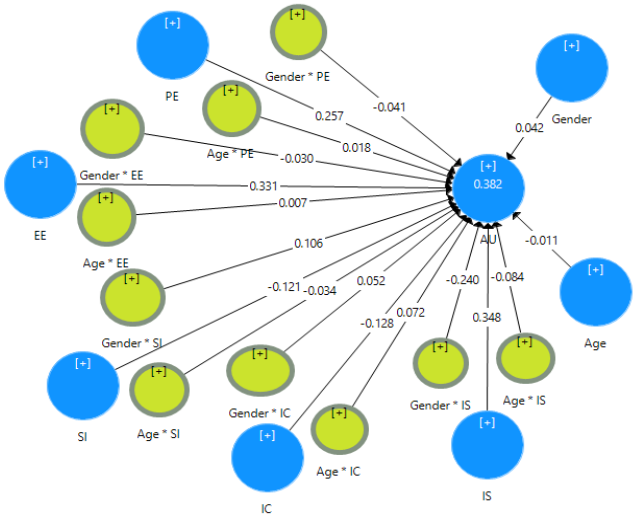


Figure 3: Structural model results

6.2. Structural Model Assessment

The bootstrap method of resampling of 5,000 iterations was performed to test the coefficients of the individual paths of the structural model. According to [39] the structural model's explanation relies on the coefficient of determination, R^2, to explain endogenous variable. The results of the structural model are given in Figure 3.

Table 4: Path Coefficients

Hypotheses	Path	Path coefficient	t-statistics	p-values	Result
H1a	PE -> AU	0.26	2.66	**0.01**	Confirmed
H1b	Gender * PE -> AU	-0.04	0.37	0.71	Unconfirmed
H1c	Age * PE -> AU	0.02	0.18	0.86	Unconfirmed
H2a	EE -> AU	0.33	3.32	**0.00**	Confirmed
H2b	Gender * EE -> AU	-0.03	0.32	0.75	Unconfirmed
H2c	Age * EE -> AU	0.01	0.07	0.95	Unconfirmed
H3a	SI -> AU	-0.12	1.31	0.19	Unconfirmed
H3b	Gender * SI -> AU	0.11	0.96	0.34	Unconfirmed
H3c	Age * SI -> AU	-0.03	0.39	0.69	Unconfirmed
H4a	IC -> AU	-0.13	1.39	0.16	Unconfirmed
H4b	Gender * IC -> AU	0.05	0.53	0.60	Unconfirmed
H4c	Age * IC -> AU	0.07	0.85	0.40	Unconfirmed
H5a	IS -> AU	0.35	3.40	**0.00**	Confirmed
H5b	Gender * IS -> AU	-0.24	2.36	**0.02**	Confirmed
H5c	Age * IS -> AU	-0.08	0.74	0.46	Unconfirmed

6.3. Hypotheses Testing

From Table 4, PE had a significant effect on actual use of LMS ($\beta = 0.26$, $p = 0.01$), hence supporting H1a. Also, the effect of effort expectancy on students' actual use of LMS was significant ($\beta = 0.33$, $p = 0.00$), supporting H2a. Furthermore, IS had significant impact on actual use of LMS ($\beta = 0.35$, $p = 0.00$), supporting H5a. However, SI ($\beta = -0.12$, $p = 0.19$) and IC ($\beta = -0.13$, $p = 0.16$) did not have a significant effect on the actual use of LMS, hence dH3a and H4a were not supported. Regarding the moderating effects of age and gender, gender moderated significantly on the correlation between institutional support and actual use of LMS ($\beta = -0.24$, $p = 0.02$). Moreover, Figure 4 shows that, as institutional support increases the actual use also increases

for both males and female, but the increment is higher in females than in males, confirming H5b. However, the hypotheses H1b, H1c, H2b, H2c, H3b, H3c, H4b, H4c, and H5c were all not supported.

The actual use of LMS was significantly explained by PE, EE, SI, IC and IS, and the moderating effects of gender and age at R^2 = 0.38, showing the five independent variables and the two moderating variables explained 38% of the change in AU. Moreover, in this work, the value for the standardized root mean square residual (SRMR) was 0.069, suggesting a good model fit. This value confirms [40] assertion that SRMR value less than 0.1 gives an acceptable fit model. The result of the structural model is shown in Figure 3.

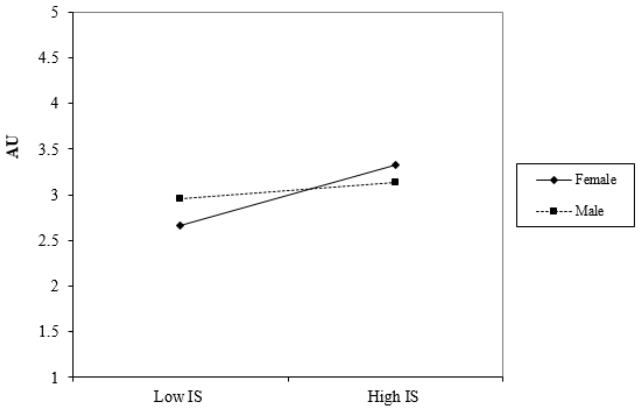


Figure 4: Moderating effect of gender on the correlation between Institutional Support (IS) and Actual Use (AU)

7. Discussions

This study intended to examine the determinants of students' actual use of the learning management system (LMS). [41] stated that knowing the determinants of an e-learning system is key to acquiring better discoveries and answers [42] also stressed that students and teachers should be encouraged to use the e-learning system. The research model accounted for 38% of the change in students' actual implementation of LMS, confirming that the new model is suitable in explaining learners' implementation of LMS.

The association between PE and actual usage of LMS was shown to be significant, implying that the greater the PE of a system, the greater the actual usage. Therefore, administrators should consider the benefits the students will gain from a system when planning to acquire one. This result was in agreement with previous studies [13], [14].

Furthermore, there was a significant correlation between EE and actual, implying that students will be willing to use the LMS if it is flexible, but will refuse to use the system if it is not flexible to use. This finding is in agreement with [11]. Additionally, this study revealed that SI did not significantly relate to actual use. This result is in disagreement with previous studies [23], [24]. This finding could be explained that the students are technology savvy and know the benefits of technology hence they do not need peers or any other entity to influence their use. Also, it was found that the IC did not impact the actual use of LMS. This means that IC singly was not sufficient to stimulate learners' use of the system.

Finally, it was revealed that IS significantly related to students' actual use of LMS. This result is in agreement with prior works such as [27] [28]. Moreover, it was reported that as the institutional support increases so as actual usage increases for both males and females, but the increment is more significant in females than in males. This is in agreement with [16] who found out that gender moderated significantly on the correlation between facilitating condition and intention to use technology. According to [29], training and support are significant elements of individuals' use of technology.

8. Conclusion

The present study contributes to knowledge by providing first-hand data that the research model elucidates the determinants of learners' technology use. The research model elucidates 38% of the changes in learners' technology use. This finding has provided empirical data for studying several determinants that impact on students' actual application of technology. From a research point of view, this study primarily attempted to validate the findings from previous studies by empirically determining learners' actual implementation of LMS using UTAUT in unindustrialized nations. Future studies should examine other determinants such as technology self-confidence, individual innovativeness, and attitude that influence the usage of technology [43]. Although these may further explain the actual implementation of technology, it cannot assure us that the coefficient of determination of a new framework would be improved. To conclude, this study found that the new model was suitable for explaining the determinants of learners' actual application of technology.

9. Implication for Practice

PE and EE were found to be strongly related to the actual use. Irrespective of the usefulness of technology, users would be unwilling to use the technology if they find it to be cumbersome and difficult to use. Hence, management should consider the user-friendliness and flexibility of technology, as well as, its usefulness when adopting one. Additionally, educating academicians and learners would promote the use of technology in the instructional process. Furthermore, management in educational institutions should pay more attention to the provision of institutional support for learners and instructors. This would encourage lecturers and learners to use the online learning system to boost their self-confidence as well. [44] concluded that institutional support has a positive effect on the use of the e-learning learning system.

10. Drawbacks and Future Studies

The drawbacks of this study are: (1) Since this study involved one university, it is difficult to generalize the finding to other institutions, considering variations in the use of technology in different institutions. Therefore, future research should be done in other universities. (2) Thirty-eight percent (38%) of the change of

the dependent variable was elucidated by the exogenous variables. Other variables should be considered in future studies to unearth the outstanding 62%. (3) In this study, a cross-sectional strategy was adopted to gather data. In future studies, the longitudinal method ought to be employed to collect data from the partakers. (4) the use of other moderating factors is recommended for future studies. Finally, instead of using a questionnaire to collect data, future studies should use interviews to collect data from the partakers.

References

[1] A. Hassanzadeh, F. Kanaani, S. Elahi, "A model for measuring e-learning systems success in universities" Expert Syst. Appl, 39 (12), 10959 –10966, 2012.

[2] K. A. Al-Busaidi, K. and H. Al-Shihi. "Instructors' Acceptance of Learning Management Systems: A Theoretical Framework," Communications of the IBIMA, vol. 2010, 2010.

[3] C.S. Nair, R. Patil, "A Study on the Impact of Learning Management Systems on Students of a University College in Sultanate of Oman", Int. J. Comput. Sci, 9(2), .379–385, 2012.

[4] R, Ahmad, R. Edwards, R. Tomkinson, "The use of WebCT in distance learning course in university of manchester", The Turkish Online J. Dist. Educ, 7(2), 101-108, 2006.

[5] C. Hong, Walker, D. A, "A confirmatory factor analysis of the Technology Acceptance", *General Linear Model J.l*, 41(2), 22-29, 2015.

[6] A.N. Islam, "Investigating e-learning system usage outcomes in the university context", Comput. Educ., 69, 387–399, 2013.

[7] M.. Waheed, K. Kaur, N. Ain, N. Hussain., "Perceived learning outcomes from moodle: an empirical study of intrinsic and extrinsic motivating factors", Inf. Dev, 32(4), 1001-1013, 2015.

[8] N.D. Oye, N.A. Iahad, N. Ab.Rahim, "The history of UTAUT model and its impact on ICT acceptance and usage by academicians", Educ. Inf. Tech, 19(1), 251-270, 2014.

[9] B. Bervell and I. N. Umar, "A decade of LMS acceptance and adoption research in Sub-Sahara African higher education: A systematic review of models, methodologies, milestones and main challenges", Eurasia J. Math. Sci. Tech. Educ, 13(11), 7269–7286, 2017.

[10] T. Teo., N. Ruangrit, J. Khlaisang, T. Thammetarand, K.S. Jumnong, "Exploring e-learning acceptance among university students in Thailand: A National Survey", J. Educ. Comput. Res, 50(4), 489-506, 2014

[11] V. Venkatesh, M.G. Morris, G.B. Davis, F.D. Davis, "User acceptance of information technology: Toward a unified view", MIS Quart, 27(3), 425–478, 2003.

[12] W.R. King, J. He, "A meta-analysis of the technology acceptance model", Inf. Manage, .43(6), 740–755, 2006

[13] N.D. Evans, and J. le Roux, "Modelling the acceptance and use of electronic learning at the University of Zululand", S. Afric. J. Librar. Inf. Sci, 81(2), 26-38, 2015.

[14] P.C. Nanchang, "Student Acceptance of Web-based Learning System.Proceeding of the 2009 International Symposium on Web information Systems and application (WISA 2009)", China, 2009.

[15] B. Sumak, G. Polancic, M. Hericko, "An empirical study of virtual learning environment adoption using UTAUT in 2010 Second Int. Conf. on Mobile, Hybrid, and On-line Learning ELML. Saint Maarten, 2010.

[16] K. Ghalandari, "The effect of performance expectancy, effort expectancy, social influence and facilitating conditions on acceptance of e-banking services in Iran: The moderating role of age and gender", Mid-East J. Scient. Res., 12(6), 801-807, 2012

[17] O. A. Aliyu, C. Arasanmi, S. Ekundayo, "Do demographic characteristics moderate the acceptance and use of the Moodle learning system among business students?", Int. J. Edu. Dev. Comm. Tech., 15(5), 165-178, 2019.

[18] Wang, Y.-S., Shih, Y.-W, "Why do people use information kiosks? Validation of the Unified Theory of Acceptance and Use of Technology", Govern. Inf. Quart, 26(1), 158-165,2009, doi:https://doi.org/10.1016/j.giq.2008.07.001

[19] K. Magsamen-Conrad, S. Upadhyaya, C.Y. Joa, J. Dowd, (2015, "Bridging the divide: Using UTAUT to predict multigenerational tablet adoption practices", Comput. Hum. Behav, 50, 186-196, 2015.

[20] N. Ain, K. Kaur, M. Waheed, "The influence of learning value on learning management system use: An extension of UTAUT2", Inf. Dev, 32(5), 1306-1321, 2016.

[21] A. Raman, Y. Don, "Preservice teachers' acceptance of learning management software: An application of the UTAUT2 Model", Int. Educ. Stud, 6(7), 157–164, 2013.

[22] T. Ali, H. Kate and L. Xiaohui, "Factors Affecting Students Acceptance of E-learning Environments in Developing Countries: A structural Equation Modeling Approach" Int J. Inf. Educ. Tech, 3(1), 54-59, 2013.

[23] Y.L. Wu, Y.H. Tao, P.C. Yang, "The use of unified theory of acceptance and use of technology to confer the behavioral model of 3G mobile telecommunication users", J. Stat. Manage. Syst, 11(5), 919-949, 2008

[24] A. Raman, Y. Don, R. Khalid, M. Rizuan, "Usage of Learning Management System (Moodle) among Postgraduate Students: UTAUT Model", *Asian Soc. Sci.*, 10(14), 186-192, 2014

[25] A. Fidani and F. Idrizi. "Investigating students' acceptance of a learning management system in university education: A structural equation modeling approach" in 2012 ICT Innovation Web Proceedings, Macedonia, 2012.

[26] V. Venkatesh, Bala, H, "Technology acceptance model 3 and a research agenda on interventions", Decis. Sci, 39(2), 273-315, 2008.

[27] M. S. H. Khan, "Barriers to the introduction of ICT into education in developing countries: The example of Bangladesh", Int. J. Inst, 5(2), 61-80, 2012.

[28] J. Brinkerhoff, "Effects of a long duration, professional development academy on technology skills, computer self-efficacy, and technology integration and beliefs", J. Res. Tech. Educ, 39(1), 22 – 43, 2006.

[29] D. Jenkins, C. Mimbs, T. Kitchel, "Computer Literacy, Access, and Use of Technology in the Family and Consumer Sciences Classroom", J. Fam. Consum. Sci. Educ., 27(1), 1–13, 2009.

[30] R. Chen, "Investigating models for preservice teachers' use of technology to support student-centered learning", Comput. Educ., 55(1), 32–42, 2010.

[31] B.M. Sharan, Qualitative Research: A Guide to Design and Implementation, San Francisco: Wiley, 2009.

[32] B. Flyvbjerg. "Five Misunderstandings About Case-Study Research", Qual. Inqui., 12(2), 219-245, 2006.

[33] J.F. Hair, G.T.M. Hult, C. Ringle, M. Sarstedt, A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), Thousand Oaks: Sage Publications, 2016.

[34] A. Suh, C. Wagner, L. Liu, "Enhancing User Engagement through Gamification", *J. Comput. Inf. Syst*, 58(3), 204-213, 2018.

[35] A. Ahmed et al., "Modeling and Simulation of Office Desk Illumination Using ZEMAX," in 2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), 2019, pp. 1–6.

[36] C. Fornell, D.F. Larcker, "Evaluating structural equation models with unobservable variables and measurement error", J. Mark. Res, 18(1), 39-50, 1981

[37] W.W. Chin, "The partial least squares approach to structural equation modeling. In G.A. Macoulides (Ed.), Methodology for business and management. Modern methods for business research", Lawrence Erlbaum Associates Publishers, 1998.

[38] J, Henseler, C.M. Ringle, Sinkovics, R.R, "The use of partial least squares path modeling in international marketing. In R.R. Sinkovics and P.N. Ghauri, Eds., New Challenges to International Marketing", Emerald Group Publishing Limited, 2009.

[39] K.S. Ofori, H. Boateng, A.F. Okoe, I. Gvozdanovic, "Examining customers' continuance intentions towards internet banking usage", Mark. Intell. Plann, 35(6), 756-773, 2017

[40] C. Sengul, E. Ilker (2015) "Comparison of model fit indices used in structural equation modeling under multivariate normality", J. Mod. Appl. Stat. Meth, 14(1), 2015.

[41] R. Boateng, A.S. Mbrokoh, L. Boateng, P.K. Senyo, "Determinants of e-learning adoption among students of developing countries", The Int. J. Inf. and Learn Technol, 33(4), 248-262, 2016

[42] D.D. Prior, J. Mazanov, D. Meacheam, G. Heaslip, J. Hanson, "Attitude, Digital literacy and self-efficacy: Flow-on effects for online learning behavior", The Internet Higher Educ., 2016.

[43] V. Venkatesh, J.Y. Thong, F.K. Chan, P.J.H., S.A. Brow, "Extending the two-stage information systems continuance model: Incorporating UTAUT predictors and the role of context", Inf. Syst. J, 21, 527–555, 2011.

[44] Z. Yildrim, C.M. Reigeluth, S. Kwon, Y. Kageto, Z. Shao, "A comparison of learning management system in a school district: Searching for the ideal personalized integrated educational system (PIES)", Interactive Learn, Environ, 22(6), 721-736, 2014

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 621-634 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Commercialization Process of Disruptive Innovations in Corporate Ventures and Spinoff Companies: A Comparison

Javier Nieto Cubero*,1, Saheed Adebayo Gbadegeshin[2], Carolina Consolación Segura[3]

[1] *EINA University Center of Design and Arts, Autonomous University of Barcelona, Design Processes Research Group, 08017, Spain*

[2]*Turku School of Economics, University of Turku, Department of Management and Entrepreneurship, 20014, Finland & DeGesh Institute of Technology and Entrepreneurship; 20002, Nigeria*

[3]*ETSETB, Polytechnic University of Catalonia, Department of Management, 08034, Spain*

ARTICLE INFO

Article history:
Received: 11 January, 2020
Accepted: 22 March, 2020
Online: 14 April, 2020

Keywords:
Commercialization
Transition Management
Disruptive Innovations
Ambidexterity
Corporate Venturing
Spin-off

ABSTRACT

One of the most critical challenges that large companies, small enterprises and research institutes face, when commercializing their innovations, is the transfer process at the moment of scaling up. These organizations often transfer their innovation to either existing business units or create a new business entirely. The process of transferring innovations to an existing unit is known as corporate venture, while creating a new business is called spinoff. Both processes have been studied separately, but not yet together and from different business sectors and countries. In order to understand both processes, this article compares the commercialization of disruptive innovations from these sectors: electronics, information technology, telecommunications, engineering, healthcare, biotechnology or research equipment. The article used multiple case study methodology from commercialization projects carried out in large, small and medium-sized companies and research institutes based in Spain, Germany, UK and Finland. The findings showed that certain activities needed to be done at pre-transition /commercialization, transition/commercialization, and post-transition/commercialization phases to reach a successful transition. Furthermore, the study provides similarities between the corporate venture and spinoff approaches that include team formation, network development and getting the commitment from the company and stakeholders. Additionally, the article outlines divergences between the approaches which consist of innovation readiness, knowledge management or the activities of the post-transition phase. The article also provides insight for innovation scholars, commercialization practitioners, and business enterprises. Therefore, the article contributes to the commercialization of disruptive innovations.

1. Introduction

In today's fast-paced technologically based environments, to keep long-term competitiveness, companies need to introduce disruptive and incremental innovations to markets [1,2]. The disruptive innovations need to be commercialized as it is argued by the scholars [3]. Unfortunately, the commercialization of these innovations has not been thoroughly studied [4], despite its importance within the innovation process [5]. Disruptive innovations are challenging to commercialize by both established enterprises and new businesses (specifically spinoffs). One of the most critical challenges that mature corporations face at the commercialization phase is the transition of an innovation project to their business units for scaling up the project. Similarly, spinoff companies meet obstacles related to organizational factors when commercializing their disruptive innovation [6].

Thus, the corporate venture could be described as the exploration and exploitation of new technologies or products. It is detached from the corporate parent to avoid organizational inelasticity. Additionally, the term "corporate venture" could refer to "spin-in" according to [7]-[9]. Moreover, the commercialization of innovations by establishing a new enterprise is known as spinoff. Spinoff is defined as the creation of a new company to

*J. Nieto Cubero, P. Santa Eulalia 25, 08017 Barcelona, +34 653560568, jnietoc@eina.cat

commercialize the business potential of a technology or innovation.

Currently, literature has explained the corporate venture process from many different perspectives, such as the corporate parent [10] or the corporate venture [11]. However, it is noted in the literature that there is a limited discussion on the venture transition at the moment of scaling up the innovation. Specifically, there is a limited scholarly discussion on how to design the transition process. Similarly, there is not yet a model that discusses activities needed for the successful transition of the corporate venture into the established company businesses [12] to reach the full commercialization of the innovation.

The existing literature also has explained the spinoff process from the parent organization and the academic entrepreneur's perspective. Nevertheless, few studies have focused on the commercialization process of the spinoff and how the innovation is transformed into a marketable product or service via the spinoff. Therefore, there is a need for in-depth knowledge of the commercialization process of spinoff [6].

Furthermore, ambidextrous literature has not yet provided a better explanation for the transferring of disruptive innovations in corporate entrepreneurship environments, by spin-in (corporate venture) or spinoff. Lastly, the current literature does not either provide a comparison among corporate ventures and spinoffs in their process of commercialization.

Therefore, the purpose of this article is to provide a better understanding of the commercialization process of disruptive innovations in the context of corporate venture and spinoff. It aims to provide critical factors and activities to perform in order to reach a successful scaling up innovation in these environments. To achieve these goals, it employed a qualitative research method.

The article also provides a comparison of the process of commercialization between corporate ventures and spinoffs in order to learn about the challenges they face in the commercialization process. The article offers both theoretical and practical knowledge to discourse on the commercialization process.

The rest of the article is organized as follows; Section 2 presents the theoretical background about the commercialization process in the corporate venture and spinoff, section 3 details the employed research method, section 4 presents the results obtained from the empirical studies, and section 5 presents discussion and conclusions. The last section presents contributions and limitations of the article as well as provides avenues to future research.

2. Theoretical Background

Disruptive innovations are new technologies, techniques or knowledge that generate market and technological breakthroughs on macro and micro levels. These innovations embody new technology that causes a new market environment or generates different customer behaviors. These innovations are affected by S-curve(s): market, technology, or both, during their introduction to the marketplace [13,14]. More specifically, disruptive innovations are radical, and they distort existing ecosystems [15]. Thus, the disruptive innovation changes the relations among firms and customers, modifies marketplace, overtakes current products and reconfigures the innovator's resources [16,17].

The corporate venture is often used to improve competitive positioning and transform corporations, their markets, and industries. It can also be used as opportunities to exploit value-creating innovations. The corporate ventures are startup-like structures that are idea-wise rooted in an established company. They are commonly used in exploration processes of disruptive innovations.

Thus, the corporate venture is, at this moment, described as the exploration and exploitation of new technologies or products, detached from the company's exploitation activities in an "incubation" cell structure that avoids organizational inelasticity and provokes resources and organizational renewal [18]-[20]. The corporate is notably essential when the innovation is still in an early phase, characterized by high uncertainties. Typically, the venture should be transferred to an existing business unit from the corporate parent creating profits within the firm. Meanwhile, the relationships between new corporate ventures and existing business units generate conflicts [21,22]. If this conflict is escalated, the incorporation of corporate venture into an existing corporate unit could be harmful to the development of the innovation and could also find hostility in the receiving team [23].

Notably, the corporate venture is different from intrapreneurship, which is described as infusing the mindset of entrepreneurs in the company's team and thus generating an innovative atmosphere within the company, as enunciated by [24]. Similarly, the corporate venture is different from the spinoff. While the corporate venture is a structurally separated unit with resources to manage innovation projects from idea to commercialization, the spinoff is an entirely new business created to develop and exploit the potential of the innovation.

The spinoff is defined as the generation of a new business to reach profits of new technology or innovation. This new business formation is attained by transferring innovation and other resources (e.g., intellectual property and human) from the originating organization. The originating organization consists of companies, universities or tertiary institutions and research institutes [25]. The spinoff also refers to a new firm that is established to commercialize business opportunities for a new technology that is developed by the research organizations [26]. Additionally, it could be a new company founded fundamentally to exploit an intellectual property that is developed by a university [27]. The spinoff is otherwise known as an academic spinoff, research-based spinoff or new technology-based firm.

2.1. Corporate Venture (CV)

This article focuses on the transition phase of disruptive innovations, and particularly on the transference of innovations to a receiving business unit of the corporate parent.

Usually, a split exists among innovation teams and existing business units, as stated by [13]. Supported by research from [19], these scholars point out that there is not a defined process or methodology to guarantee the delivery of the innovation projects from research and development (R&D) at an acceptable level to the business unit before reaching the scaling-up phase. It has been

studied in literature through different perspectives, and it is described in the following paragraphs.

Uncertainties perspective: the scholars [28,13] outlined four kinds of uncertainties: technical, market-concerned, organizational and resources-based. Due to these uncertainties, disruptive innovations are usually substituted by incremental innovations [29]. Technical uncertainty includes the scientific knowledge base, technical feasibility and manufacturing process. The market-concerned uncertainty comprises the understanding of the new customer value, in relation to the existing competition [13]. The organizational uncertainty involves the organizational and managerial difficulties at integrating a disruptive innovation while pursuing exploitation activities. This uncertainty is mainly caused by unsupportive top management [10] or by the misfit among the corporate business unit and the innovation team [16]. The last uncertainty is resource-based. It addresses all the issues of acquisition of the resources to get successful exploitation of the disruptive innovation [30,31].

Technology readiness: the research from [32] offers an assessment tool to convert a technology invention into a profitable business, and the study of [33] provides a global view about the Technology Readiness Levels (TRL's). Nonetheless, these scholars only approach the phenomenon under a technology uncertainty perspective.

Ambidexterity perspective: the ambidexterity literature [20] introduces the construct of separated units for exploration and commercialization, coordinated at the top management level. This structural detachment guarantees that the exploration activities are preserved from the bureaucracies, managerial routines, and the culture of exploitation [34]. The works from [35,36,20] have enunciated that the detachment among exploration and exploitation activities foster the benefits of both [37]. Similarly, the work of [36] argues that ambidextrous organizations require methodologies to activate, organize and incorporate separated exploration and exploitation units in organizations.

Separation is mainly required to emphasize on disruptive exploration. Nevertheless, the detachment of these units from the corporate parent is a necessary yet insufficient condition for ambidexterity [38].

Notwithstanding the acknowledged obstacles for venture transition, the extant literature on the process of transference to an existing business is relatively scarce [39]. Research on integration mechanisms is still needed, and as [36] underlined, is required more research about how "transition should be done."

Championship perspective: many different scholars enunciate the relevance of having the support of a champion not only in the innovation team but also at senior management level [40,41]. Several studies also have studied the role of senior management in fostering integration and enhance synergies between exploration and exploitation units [12,35]. Moreover, the likelihood of transition success is also raised if a champion appears in the receiving unit [12, 13] or if it has the support of an informal network [31]. Clear and transparent communication is critical to recognize the right supporters within the corporate parent [29]. [11] also stresses the relevance of exchanges within corporations not only at a management level but also between business units and other stakeholders of the firm.

Regarding CV, the disruptive innovation cell's managers are needed to smooth the integration of the project into the business unit by demonstrating the feasibility of the innovation, making concepts and benefits tangible and finding latent needs from customers.

2.1.1. Corporate ventures transition phases

The work from [12] distinguishes three stages in the transition from a separated venture to a one that is integrated into the established corporate structure: the pre-transition, the transition, and the post-transition phase. The pre-transition phase is described as a stage to develop connecting and learning activities between the CV and the parent business. During the transition phase, the hand-over moment takes place, and the venture is transferred to the corporate organization, while the post-transition phase is defined as the stage when the venture is incorporated in the established organization.

2.1.2. Specific activities in transition phases in corporate ventures

A summary of the existing literature on activities to perform in the commercialization process and activities to be performed within each phase of transition is displayed in Table 1.

2.2. Spinoff (CSO)

There are two approaches to the spinoff. The first approach is a planned spin-off. The planned spin-off is predetermined and is based on the plan of a parent organization. The parent organization often states that a new venture would stem from the commercialization activities of their project. The second approach is a spontaneously occurring spinoff. It happens when a new business venture is created unexpectedly. Generally, this spin-off does not get total support from the parent organization [42]. Commercializing innovation through a spinoff has been studied through different approaches. They are described in the following paragraphs.

Strategic perspective: From the strategic viewpoint, the scholars state that commercializing an innovation requires different strategies, orientations, and decisions. According to [43], there are three strategies: early adoption, adoption network configuration, and mainstream adoption. These authors explained that early adoption strategy is the method when a commercializing team adopts to create awareness, stimulate and motivate people to have a positive attitude towards the new product. Similarly, the authors stated that the adoption of a network configuration is the ability of the team to enable early buyers to be strong supporters of the product. Likewise, the authors mentioned that the mainstream strategy is the decision and effort of the team to transfer a configured network to the primary market. Still, from the strategic perspective, commercializing an innovation also requires different decisions to be made. [44] found that world-leading companies, well-known for their innovation commercialization, made critical decisions by placing commercialization as one of their highest priorities, establishing measurable objectives for proper management, and developing sufficient and essential competencies.

Table 1. Corporate venture activities in commercialization phases

Transition Phase	Category	Specific Activities	Scholars
Pre-Transition	Innovation project preparation for scale-up	Transition readiness assessment	O'Connor, Hendricks and Rice (2002), Rice, Leifer and O'Connor (2002) and van Burg et al. (2012).
		Transition team	Rice, Leifer and O'Connor (2002) and van Burg et al. (2012).
		Transition plan	Kanter et al. (1991) Rice, Leifer and O'Connor (2002).
	Conditions for validation	Achievement of first sales and running of pilot projects	Nieto (2018).
	Early-relations with the potential receiving business units	Training and personal exchange	van Burg et al. (2012).
		Cross-functional interfaces	Jansen et al. (2009).
		Liaison channeling and network building	Gassman et al. (2012).
		Horizontal interactions between teams	Raisch and Tushman (2016).
Transition	Transition time	Transference after achieving first sales	van Burg et al. (2012).
		Laying the groundwork for a significant market	Rice, Leifer and O'Connor (2002).
	Knowledge management	Transfer of the R&D team to the receiving business unit	Nieto (2018).
Post-transition	Specification of suitable KPI's in the business unit	Post transition autonomy. Performance metrics	van Burg et al. (2012).

Marketing perspective: The scholars, from the marketing viewpoint, seem to focus on the marketization of the disruptive innovation through innovation, according to [2]. The high technologies make use of disruptive innovations. Thus, marketing scholars state that innovation would be successful if the final consumers can accept them. Examples of prominent scholars in this group are [45,46]. To reinforce the credibility of this viewpoint, [47] argued that innovation would be successfully commercialized if it could be market oriented. This scholar stated that commercialization must start with the market and end with it. In a similar view, [48] added that the commercialization should focus on the substitute products/services as well as on the functionality of commercializing products concerning its price. However, marketization is cumbersome, tedious and full of risks. According to [49], six challenges are awaiting the commercializers of disruptive innovation. The scholar stated that they market-concerned uncertainty, technological uncertainty, the inconsistency of new products in a complex multi-component system, difficulty in developing networks, the problem of ecosystem complexities and competition, and inherent risks of choice-making (especially with multiple and interdependent product-market options).

Skills perspective: From the skills perspective, the scholars emphasize how the commercialization should be accomplished. Examples of scholars in this perspective are [50,51]. [51] emphasized that innovation can reach the market successfully if the commercialization team can: (1) acquire and possess sufficient technological knowledge, (2) develop the innovation into a product which can be mass-produced, (3) establish commercialization process and relate it to growth strategy, (4) adhere to their process as a learning path, and coordinate and maintain a good network. Furthermore, [52] stressed that, though focused on the strategy, marketing skills are crucial for their successful commercialization. To assist practitioners, these scholars proposed market and user research tools, which include customer ethnography, emphatic designs, lead-user processes, investigating users, and the targeting of new markets.

Considering all these perspectives, the question of "what one has to do in commercializing disruptive innovation successfully in terms of activities" is fragmented depending on the perspective approach.

2.2.1. Spinoff commercialization phases

When commercializing disruptive innovations through a spinoff, three phases are recognized, as acknowledged by [53]-[56]. These phases are pre-commercialization, commercialization and post-commercialization. The pre-commercialization is the stage where any activity aiming at transforming new technology into products and services is done. The commercialization phase contains all the transformation efforts. The post-commercialization includes any attempt to make new products or services sustain market share at a profit for its commercializing team and companies.

2.2.2. Specific activities in commercialization phases in spinoffs

Based on the above grouping, the previous works describe the activities to perform in each phase and they are summarized in the following Table 2.

2.3. Relationship between CV's and CSO's in the commercialization stage.

The work from [57] discussed the role and relevance of internal corporate venturing and spinoffs as means to corporate renewal and to improve competitive positioning.

Meanwhile, similarities and differences between different corporate entrepreneurship structures have been researched by scholars, for instance between Startups and CV's, as enunciated by

Table 2. Spin-off activities in commercialization phases

Commercialization Phase	Specific Activities	Scholars
Pre-commercialization	Basic and applied science /discovery technology/ initiation /	Maine and Garnsey (2007), AbdRahima et al. (2015), Amadi-Echendu and John, (2008), and Pietzsch et al. (2009).
	Formulation/identification of technological opportunity/ Application idea initial technical and economic viability	Maine and Garnsey (2007), AbdRahima et al. (2015), Amadi-Echendu and John (2008).
	Reviewing the technology / Application	Eldred and McGrath (1997) and Amadi-Echendu and John (2008).
On-commercialization	Having a development team and organizing a senior review team / formalizing commercialization project	Eldred and McGrath (1997) and Rogers et al. (2004).
	Formalizing commercialization project	Rogers et al. (2004).
	Scanning and creation of potential alliances/identification of specific need of target market/ conducting of preliminary material investigation /	Maine and Garnsey (2007) and Chen and Panda (2005).
	Designing, building and testing of prototypes/designing and prototyping/design and development / Developing prototypes and integrating the prototype into existing products/incubation	Rogers et al. (2004), Maine and Garnsey (2007), Pietzsch et al. (2009) and Chen and Panda (2005).
	Development of technological product/service	AbdRahima et al. (2015) and Amadi-Echendu and John (2008).
	Customer testing and experimentation	Maine and Garnsey (2007).
	Verification of relating policies/ evaluation of make/buy decision/development of the pilot plant,	Rogers et al. (2004) and Maine and Garnsey (2007).
	Deciding on channels of distribution	Rogers et al. (2004).
	Final validation / final customer testing	Pietzsch et al. (2009) and Maine and Garnsey (2007).
	Developing a structured commercialization method / developing structured processes	Eldred and McGrath (1997) and Chen and Panda (2005).
Post-commercialization	Product launching and post-launching assessment/market, and acceleration	Pietzsch et al. (2009), Amadi-Echendu and John (2008).

[58], USO's (university spinoffs) and CV's, as stressed by [59], USO's and CSO's, as outlined by [60,61], or between IV's (independent ventures) and CV's [62]. Similarly, under the perspective of the leaders of the innovation units, commonalities and divergencies between corporate entrepreneurs and spinoff entrepreneurs have been discussed by [63].

The previous works show that the commercialization process of the corporate venture consists of pre-, transition, and post-transition. The phases are similar to the spinoff, which include pre-commercialization, commercialization and post-commercialization. The previous literature also states the different activities to be done at each phase. However, there is missing knowledge of the commercialization process of disruptive innovation via the corporate venture and spinoff. Thus, this article seeks to provide to the following research questions:

- What are the key activities that facilitate managing the transition and commercialization of disruptive innovations in corporate ventures and spinoffs?
- What are the differences and similarities in the process of commercialization between a corporate venture and a spinoff?
- Do corporate ventures face the same problems that spin-offs are confronted with?

3. Research design

3.1. Methodology approach

In order to answer the research questions, a multiple case study approach has been selected. Case study research involves the examination of a phenomenon in its natural environment. The case study method is specifically appropriate to research new field areas, with a focus on "how" or "why" questions, related to a contemporary set of events. The study of multiple cases is usually considered as a more robust method, as it provides the observation and analysis of a situation in different settings. The multi-case method enables an understanding of the phenomenon beyond each project context and increases generalizability [64]. There were two sets of empirical data in this research. The first set was gathered for corporate ventures and the second for spinoffs.

3.1.1. Corporate venture case selection and data collection

Potential disruptive innovation projects were evaluated according to criteria provided by the definitions of [65], who uses disruptive to describe innovations that could occur in a macro-level as well as micro-level contributions.

Based on the above criteria and according to [66], who state that theoretical sampling is a means to reach a high gain of insight, twelve corporate venture transition processes were selected. Nine of those samples found an internal existing business unit of the corporate parent to be integrated in. One of them found an external firm, and two projects employed a dual destination model to an internal business unit as well as to an external company. Eight of the samples have been successful in the transition process, and the other four were failure cases. The cases are presented in Table 3.

The companies have been named (Telco Co, FMCG Co, Appliances Co, Financial Co, Engineering Co). We also define the

name of the chosen projects as Tel1, Tel2, Tel3, Tel4, Tel5, Tel6, Fast1, App1, App2, Fin1, Eng 1, and Eng2. The real names of companies and projects are not displayed due to the confidentiality research agreements.

For the sample, data were collected through interviews with innovation unit managers and corporate innovation managers to achieve different perspectives on the success of the project, conflicts, key issues and evolution of the project transition, reaching a variety of insights. The research carried out 26 double interviews. Each interview took 1 to 1,30 hrs. Long. The interviews were performed in two rounds. The first round, to understand the overall context of the project and company and the second round to obtain a deeper understanding of each project's details. During the conduction of the interviews, nevertheless, the interview guideline was regularly updated and enriched to be adapted to the insights obtained after the previously analyzed project. The documentation was complemented with secondary data about each project obtained from the corporate company.

3.1.2. Spinoff case selection and data collection

According to the criteria mentioned above, twenty spin-off companies were selected. They were chosen according to the type of disruptive innovations. Also, these companies were founded by serial entrepreneurs and business advisors. The details of the projects are also presented in Table 4. The spin-off projects were denoted by COY 1 to COY 20, respectively. The cases are presented in Table 4. Data were obtained through interviews. The study participants were contacted via email and telephone calls. After booking a date and venue, they were sent interview questions with themes. After a few weeks, the interviews were conducted. Each interview took more than 1 hour. All interview processes were following the qualitative research guideline provided by [67]-[71].

3.2. Data analysis

Qualitative analysis needs a different approach than the quantitative analysis because of the nature of the collected data, mostly textual and descriptive. The research is focused on the process of transition and commercialization, and more specifically, on the tasks developed to overcome the challenges found during the process. The information obtained from interviews was transcripted and later analyzed by using open coding [72] to recognize patterns in transition activities and critical problems influencing the transference and commercialization of the innovation projects. The following step was the axial coding, identifying relationships between categories, organizing them hierarchically with interconnections and sub-categories. The last step was selective coding to produce a theory and recognize core categories. NVivo11 software was used for helping in the analysis and codification of qualitative data obtained at interviews. It is a software that supports data management, in the process of exploration and coding that is not linear, but iterative.

Similar patterns responding to the formulated research questions were codified, defining the main tasks to perform in order to overcome commercialization problems and barriers in corporate ventures and spinoffs. The theory that emerged from the interviews was compared with the existing literature leading to specific conclusions displayed in the following sections. Discovered similarities increase the validity of the findings and link old and new theory. Emerging differences show opportunities for new concepts and avenues for future research.

Table 3. Summary of corporate venture cases and interviewed profiles

Parent Company	Industry	Corporate Venture	Interviewed Profiles
TELCO CO	Telecommunications	Corporate organization	Product innovation director Design Research Lead
		Tel1 project	Head of product/project Head of Innovation Portfolio Innovation Business Development Manager
		Tel 2 project	Corporate Venture Leader Head of Innovation Portfolio
		Tel 3 project	Senior Technical Expert Head of Innovation Portfolio
		Tel 4 project	Service innovation Lead Head of Innovation Portfolio
		Tel 5 project	Head of commercial innovation Head of Innovation Portfolio Innovation Business Development Manager
		Tel 6 project	Corporate Venture Leader Head of Innovation Portfolio
FMCG CO	Laundry & Home Care	Fast 1 project	Innovation Manager - IoT Corporate Venture Leader
APPLIANCES CO	Home appliances	App 1 project	Innovation Transference Director Corporate Venture Leader
		App 2 project	Innovation Transference Director Corporate Venture Leader
FINANCIAL CO	Financial	Fin 1 project	Service Design Lead Product Innovation Director
ENGINEERING CO	Engineering	Eng 1 project Eng 2 project	Division Manager Division Manager

Table 4. Summary of spinoff cases and interviewed profiles

Companies	Industry	Interviewed profiles
Coy 1	Electronics	CEO Chief Engineer
Coy 2	Environment Tech	CEO Chief Engineer
Coy 3	Healthcare Equipment	CEO Head of Venture Capital IPR Manager
Coy 4	Research Equipment	CEO Business developer Head of Venture Capital
Coy 5	Research Equipment	CEO
Coy 6	Electronics	CEO Chief Engineer
Coy 7	Research Instrument	CEO Business developer
Coy 8	Information Technology	Head of Venture Capital IPR Manager
Coy 9	Research Instrument	Principal lecturer
Coy 10	Optics	Head of Venture Capital
Coy 11	Biotechnology	Head of Venture Capital
Coy 12	Biotechnology	Head of Venture Capital
Coy 13	Electronics	Head of Venture Capital IPR Manager
Coy 14	Chemical Production	CEO
Coy 15	Information Technology	Head of Venture Capital IPR Manager
Coy 16	Electronics	Head of Venture Capital IPR Manager
Coy 17	Research Instrument	Head of Venture Capital IPR Manager
Coy 18	Information Technology	CEO Head of Venture Capital IPR Manager
Coy 19	Information Technology	Head of Venture Capital IPR Manager
Coy 20	Paper Production	Head of Venture Capital IPR Manager

4. Empirical results

The first result of the empirical study is related to the key activities performed within the different phases of commercialization. These results are explained separately for Corporate Ventures and Spin-offs in the following subsections:

4.1. Corporate ventures

Analyzing the readiness of the innovation projects, in all cases from the empirical setting, apart from having a clear value proposition for customers and stakeholders, it was also necessary the achievement of first sales to obtain the interest of potential existing corporate units. This fact was enhanced when the innovation team had conducted real pilot programs.

The critical problem to overcome, as discovered across all cases, is the misfit between the R&D project and the established corporate unit. The structure of the business unit was not prepared for the disruptive project that they received. The misfit has been discussed previously by scholars such as [13]. Also, exploitation units could lack crucial knowledge and resources to develop the innovation project. Collaboration from the business members and identifying an owner in the business unit who support and promote the project becomes one of the main challenges in full commercialization.

Another critical barrier for the survival of the transferred venture was the sales expectations the receiving business unit had on the innovation project and the lack of specific and flexible performance metrics to control the development of the venture, as previously discussed by [12].

The interviews and the data collected from the sample, allow us to identify different team activities which are crucial within the different phases of the transference of the projects to the exploitation business units, as well as the critical organizational factors influencing commercialization, providing generalization to the previous work of [73]. They are described as follows.

4.1.1. Pre-transition phase

a) Achievement of first sales and running of pilot projects.

Identifying leading customers willing to pay for the innovation is crucial in the maturity of the project. Another relevant key factor is obtaining real user traction, enhancing the interest of the corporate units, as enunciated by [73].

Moreover, the likelihood of conducting real pilot programs, and if possible, promoted by the customer enables the possibility of testing the hypotheses related to the projects linked to the particularities of the target market.

b) Early-relation with the potential receiving business unit.

Generating a relationship with the potential business unit and exchanging information concerning the R&D project seems to be crucial. The interaction between teams enables a better alignment with the corporate road map and enhances the involvement of the receiving team, as enunciated by [12,29,36,39].

It becomes relevant to recognize potential marketers in the first stages of the R&D process. These interactions could occur at managerial levels but also between the components of the teams.

This kind of relationship allows us to explain the value and potential of the innovation to gain the receiving team's trust. Communication should be transparent, transmitting the full knowledge obtained during the development phase, regarding the product, information about users and markets, including how the product has evolved and why.

Fostering communication between teams and the potential stakeholders for the innovation commercialization allows us to recognize possible future frictions, obtaining essential information to foresee how to overcome these barriers.

Finally, these cross-communications allow us to get more significant credibility for the innovation and future projects. It serves to identify key figures to obtain support from the base of the business unit, and from the management level, reaching a smoother integration.

c) Search for champions and project support network.

It has been observed in all the projects in the sample that senior management has played a supporting role, facilitating contacts and establishing bridges to look for potential marketers for the innovations. Identifying and involving a champion at the management level in the receiving team has been crucial to the success of the project. The role of informal networks has also been fundamental.

This championship perspective has been approached from the senior management perspective by [36,41]. [12,13] also pointed out the relevance of identifying an owner champion in the receiving business unit.

4.1.2. Transition phase

a) Definition of the transition time.

The different innovation unit managers underlined several tasks to get done to define the moment of transition:

- Once the R&D project has reached the first sales with a customer, it is convenient to build customer loyalty and obtain new sales with similar clients to confirm the potential of the innovation through establishing a broader and more stable customer base.
- Another crucial indicator is to have a refined and tested business and operating model ready to scale, a reliable team, and a sales road map.
- Apart from having the product ready, it is also necessary to get ready a technology platform and the complementary services to run the innovation.

These activities have not been discussed in previous literature.

b) Transfer of the R&D team to the receiving business unit

In the projects of the sample in which part of the innovation team was transferred to transmit all the knowledge acquired during the research and development phases, a positive effect was observed in the transition. This specific activity has not been discussed previously by scholars.

4.1.3. Post-transition phase

a) Performance metrics.

Business units that incorporate innovations from corporate ventures must also exploit more incremental or mature projects. For this reason, specific KPI´s must be established to apply to both the business unit teams and the disruptive innovation projects they will exploit. This concept has been previously enunciated by [12].

4.1.4. Other auxiliary tasks to carry out

Besides the similar activities identified in the projects, other patterns also appeared, described here as follows:

- Several unit managers highlighted the importance of identifying and incorporating different profiles from the business units to compose the needed team in the scaling up of the innovation. The search for these profiles aims not only to obtain the primary resources for growth but also to transform these future team members into potential champions of the innovation.
- Locating middle management champions at intermediate levels between the innovation team and the potential business unit was positive, acting as interfaces and transmitters of the benefits of the innovation.
- Another of the strategies recommended by different units and corporate innovation managers was to promote cross-movements of teams, incorporating temporary members from the innovation team into business units, enabling these profiles to approach the project from a business perspective, to detect possible future frictions.

These insights are new to literature.

4.2. Spinoff

It was noted during the interviews that technology entrepreneurs and business advisors acknowledged that certain activities have to be done before the actual commercialization and after it, defining pre-, on- and post-commercialization phases. Each phase has key activities; some of them previously not outlined in the scholarly works. The following section explains each of these critical activities according to the phases of commercialization.

4.2.1. Pre-Commercialization Phase

a) Identification of innovation application.

The study participants were able to identify an industrial process or sector where the innovation could be applied. They noted that the application should aim to solve a technical problem or social specific problem because its identification always enables commercializers to identify target markets for their innovations. This critical activity is also discussed by scholars such as [74]-[76].

b) Professional protection of Intellectual Property.

It is learned that IP protection happened after high-tech identified a business potential. The participants highlighted the importance of IP protection. They did not focus on ordinary protection, but rather professional protection. [54] emphasized the need for it.

c) Establishing of the technology - business team.

Previous works like [77,78] have stated that team formation plays an essential role in technology-based companies. The study participants plainly stated that a team should consist of someone who has a profound knowledge of the technology to be commercialized and someone who knows how to sell.

d) Defining a clear target market.

Scholars such as [45]-[47] have discussed this activity and concluded that marketization is a crucial pillar in commercialization. In the same perspective, the study participants shared that they succeeded in the commercialization because they were able to find a clear target market from the beginning of the commercialization challenge.

e) Gathering sufficient market information.

The study participants tried to gather sufficient market information. Participants stated that conducting market research ab initio is an essential step for commercialization success. They advised that commercializers should try to employ online and offline means to get market-related information on their target market. They added that the use of interpersonal networks to collect target market information played a significant role. This finding is in line with the recommendations of [48,53,57].

f) Staying with the parent company.

Participants shared that they preferred to stay with their research institutions while developing their technologies (during their early businesses). Interviewees said that it provides an opportunity to bring existing networks and resources into play during pre-commercialization. [53] acknowledged this activity in their previous study.

g) Identification of different business models.

Scholars like [79]-[81] explained that business model development is essential. This empirical study found it to be one of the tools that helped to make a business successful. Interviewees pointed out that every commercializer should know and develop several business models in which the best options can be selected in the future.

4.2.2. On - Commercialization Phase

a) Product simplification.

Participants stressed that product simplification is a key factor in the on-commercialization phase, suggesting that a new product should be able to work with existing systems or use existing infrastructure. This activity has a relationship with the recommendations of [43,48,49,77,78].

b) Subcontracting / outsourcing.

It was observed that the study participants preferred to work on their core technology and software aspects of the product while hardware and other parts were outsourced. They said that such a strategy enabled them to reach the market at the right time and to scale-up their production at a later stage of their company growth. This activity was also echoed by [43,53].

c) Having direct contact with end-users.

The participants understood that the real information for validating and improving the product solution came from end-users. Therefore, this key activity is often done during commercialization, as it is stated by [43,53,56,77,78].

d) Testing and defining flexible business models.

It is learned that the participants focused on testing while developing new business models.

e) International awareness.

The participants tried to attend international business events when developing their business models in order to improve awareness and develop networks. This activity has not yet been discussed by the previous scholars on the commercialization of innovations.

f) Reference sales.

The entrepreneur participants affirmed that one of their tasks during the commercialization stage is to make initial sales. They explained that this kind of sales is a pathway to sustainability. They shared that having references helps them to provide feedback for further development as well as create credibility for other clients and fundraising. This activity was discussed by [73] in the context of corporate ventures.

g) Starting early internationalization.

Study participants emphasized early internationalization during commercialization in order to develop a potential sales network.

4.2.3. Post-Commercialization Phase

a) Getting feedback from initial buyers

The study participants stressed that getting feedback is crucial for successful commercialization as a manner to satisfy customers as well as to offer more services, which in turn led to more revenue for the company. This activity is also outlined by [51,56].

b) Managing the supply chain

As it can be noted from the pre- and on-commercialization stages, the study participants advised commercializers to subcontract or outsource the manufacturing process. This activity is related to the work of [55], where the commercialization process from the supply chain point of view is explained.

5. Discussion and conclusions

The case studies inform us about the different activities that corporate ventures and spin-offs carry out in order to confront the main challenges they encounter at pre-commercialization, on-commercialization and post-commercialization stages. The comparative analysis between the specific activities undertaken during the phases of commercialization allows us to define the main similarities and differences among the two structures upon scaling up the innovation project.

5.1. Pre-transition/commercialization phase.

Based on the analysis of collected data, the *commitment of organizational leaders and stakeholders* to commercialize the innovation appeared to be one of the most critical factors in both corporate ventures and spinoff cases. According to [6,77] for an innovation to be commercialized successfully, the organizational leaders and stakeholders must be committed to the process. The cases of this paper revealed that upper management support motivated the commitment of teams and project managers in commercialization teams. In the case of corporate ventures, early relations with the potential receiving business unit were crucial to search owners for the innovation project, as well as to get the involvement and support from all the layers in the commercialization unit. This network building could be obtained through bidirectional personnel exchange, education and training, cross-functional interfaces or horizontal interactions between teams. Endorsing the work of [8], the empirical results suggest that these activities contribute to balance the misfit (commitment, resources and objectives) among the R&D project and the existing

business unit, a key factor influencing the successful commercialization process.

Team formation at the moment of scaling up the innovation plays an essential role in technology-based companies, corroborating the work of [77]. The specific competencies and skills required for successful commercialization encourage corporate ventures and spin-offs to define and to form the right team to face this challenge with the appropriate resources of personnel. It was noted that a commercialization team consists of people from R&D and business people, or even teams with dual technical and business development competencies. Furthermore, the project managers of these cases have the attributes of an intrapreneur, which is a closely related arrangement to the champion approach in corporate ventures. These managers can integrate technological knowledge, business expertise and user perspective. The same feature is evident among the serial entrepreneurs of the spin-off companies.

Team skills and work methodologies seem to be similar in both cases. The disruptive innovation team integrates cross profiles beyond technological qualifications. Customer research, communication and sales skills are the most valued competencies.

Collaboration or industrial networks also appeared to be another relevant factor for the transition of disruptive innovations from the parent organization to a spin-in unit or spin-off company. The cases of this paper showed that the availability of partners in and out company is a must to facilitate commercialization. It was learned that such collaboration motivated the spin-in unit to be active and it encouraged spinoff entrepreneurs to exert more effort.

Innovation project readiness. Besides these similarities, some divergences emerged from the cases studied. The innovation project seems to be more developed in the case of corporate ventures than in the case of spinoffs. Whereas in corporate ventures, the focus is on validating, demonstrating and creating credibility on the potential receiving business unit, spinoff tries to ascertain the real value or business potential of the innovation project by way of clarifying the target market. Spinoffs do so by gathering enough market information or developing several business models in which the best options can be selected in the future. Based on the collected data, in the corporate venture cases, the realization of first sales is a necessary, yet insufficient, condition for transference. It is required also to conduct pilot programs to demonstrate real traction and to maximize learning from real market situations.

This difference could be motivated by a more goal-minded orientation on the part of the project leader and team components of the corporate venture, even if they work separately from the exploitatory units of the parent company.

IP property. The analysis of the cases also reveals that corporate ventures are more financially supported and intellectual property protected thanks to their synergies with the corporate parent, while spin-offs must make an extra effort in IP protection and in search of funding and external partners.

5.2. Transition/Commercialization phase

Knowledge transference. In corporate venture cases, it is common that no one from the innovation team continues on the receiving business unit that will exploit the innovation project. Is for this reason that the study outlines the critical importance of making the right transmission of the complete knowledge acquired and of balancing out the shortcomings in terms of responsibilities within the receiving team, as well as their different motivations and goals, which confirms the previous work from [73], focused on telco industry. The information collected during the project development stage should always be transmitted to the exploitation team for a better understanding of the real needs of the users, the customer's feedbacks, the stakeholder's key information and the integration of all this knowledge into the innovation commercialization. The transference of members of the R&D team helps to manage this knowledge.

Innovation project readiness. It can be noticed that some of the activities that spinoffs perform at the commercialization phase, such as product simplification, prototyping and testing with end-users or testing business models, are a "must" type of tasks that corporate ventures have to accomplish just before the transition time. For example, the innovation should have renewed the trust not only from users but also from clients, as well as have jumped to parallel market niches. The innovation project should also have a tested and refined business and operating model, a "go to market" strategy and the associated services developed to implement the innovation.

Resources. The limited resources that spin-offs have in comparison with corporate ventures make the former focus on their core business activities using subcontracting, while the latter devote their primary efforts to finding the right resources of personnel to exploit the disruptive innovation project.

Awareness and networks. Some of the activities that spin-offs perform at this phase are oriented to maximize awareness and to grow the sales network by the participation in business events, early internalization, creation of potential alliances and the definition and search for the right distribution channels. The synergies with the parent company in the case of corporate ventures help significantly when it comes to facing these challenges.

5.3. Post-transition/commercialization phase

It is during this phase when the approach from the two structures seems to be completely different, according to the analysis of the data collected from corporate ventures and spin-offs.

The following Table 5 summarizes the comparison between activities carried out by corporate ventures and spin-offs and the main challenges to overcome during the different phases of commercialization.

The insights gained from the spinoff cases reveal that even at this phase of post-commercialization, the process of refinement of the product is crucial, requiring the feedback from initial buyers.

Even if continuous refinement is also present in the mindset of corporate ventures, the activities performed by spinoffs during this phase are usually carried out by corporate ventures at the pre-transition phase so that they can offer an optimized product to the receiving business unit.

Table 5. Comparison between corporate venture and spin-off activities and challenges during the different phases of commercialization.

	SPINOFF		CORPORATE VENTURE		
PHASE	CHALLENGES TO OVERCOME	ACTIVITIES	ACTIVITIES	CHALLENGES TO OVERCOME	PHASE
PRE-COMMERCIALIZATION	GETTING COMMITMENT FROM STAKEHOLDERS	Identification of technological opportunity Application. Technical and economic viability	Early relations with potential business units Personal exchange and training Cross-functional interfaces	COMMITMENT FROM BU. SEARCH FOR CHAMPIONS	PRE-TRANSITION
	ASCERTAINING INNOVATION VALUE AND BUSINESS POTENTIAL	Targetting markets Formulate different business models Gathering enough market information	Achieving first sales and running pilot programs	VALIDATION AND DEMONSTRATE REAL TRACTION	
			Reduce all remaining uncertainties	INNOVATION PROJECT READINESS	
	TEAM LACK OF RESOURCES AND SKILLS	Formation of the commercialization team	Definition and formation of the right team	TEAM LACK OF RESOURCES AND SKILLS	
	PROTECTION OF INNOVATION	IP protection			
ON-COMMERCIALIZATION	CONDITIONS FOR COMMERCIALIZATION	Reference sales Product simplification. Ready or existing Technology platform Testing business models Prototype and testing with end-users. Final validation	Renew sales and open new niches Clear operating model and technology platform technology platform. Complementary services Ready business model	CONDITIONS FOR TRANSITION TIMING	TRANSITION
	LACK OF RESOURCES	Subcontracting	Search for the right personal resources	RESOURCES LEVERAGING	
	AWARENESS AND SALES NETWORK	Participation in business events Internationalization Creation of potential alliances Searching for distribution channels	Synergies with corporate parent	AWARENESS AND SALES NETWORK	
			Team transference for knowledge transferences	KNOWLEDGE MANAGEMENT	
POST-COMMERCIALIZ.	PROMOTING SPINOFF	Product launching and post launching. Market assessment			POST-TRANSITION
	REFINING INNOVATION PROJECT	Getting back from initial buyers			
	OPTIMIZATION VALUE CHAIN	Managing supply chain			
			Defining specific performance metrics	BALANCE EXPLORATION & EXPLOITATION PROJECTS	
			Searching for partial autonomy	KEEPING CV MINDSET	

The size and lack of resources that spinoffs have, push these kinds of organizations to focus part of their specific activities on this phase, in order to define and develop the supply chain, the commercialization, and the launching process that will help to scale up the innovation.

The study shows that in the cases of innovation projects from corporate ventures transferred to a business unit, they focus on keeping some autonomy of criteria and mindset (coming from the R&D team) in order to optimize the scaling up of the innovation in the corporate parent. Also, participants asked to define specific performance metrics that could be applied to the commercialization team in order to be able to balance explorative and exploitative projects.

6. Contributions and limitations

The main contribution of this article is that it provides the key commercialization activities for the corporate venture and spinoff. It outlines that team formation, network development, and early relations with potential comercializers and stakeholders are the important commercialization activities that the disruptive innovations require regardless the commercialization approach. Additionally, the article outlines that there are divergences between the corporate venture and the spinoff approach. These divergences include innovation readiness, knowledge management or the approach to the post-transition phase.

Another contribution of the article is the comparison of both corporate ventures and spinoff. Even though there are many literary works on both approaches, there is no scholarly work yet that compares them. This comparison shows that both approaches have various phases and each of the phases has several activities. It also provides an in-depth understanding of both approaches on the commercialization of disruptive innovation, stressing the common challenges corporate venture and spinoff face, as follows: ascertaining and validating the business potential of the innovation, getting the commitment from the corporate parent and stakeholders, overcoming the lack of resources and obtaining awareness and credibility from market and stakeholders.

Additionally, the article contributes to the practice by enhancing understanding of the commercialization process of disruptive innovation via the corporate ventures and spinoffs approaches. It outlines some key barriers that the practitioner might face during the process, as the misfit between the R&D project and the established corporate unit, the lack of specific and flexible performance metrics to control the evolution of the project or how to define the right conditions for the commercialization of the innovation

However, the article has some limitations, defining at the same time fields for future research. Firstly, the case study method applied in the article inherited the problem of generalization, as it is usually the case. This makes the findings of this article to be limited with regards to generalization. Although the findings are generic, which means they can be generalized, yet it is recommended that the findings' generalization is determined by the nature of the concerned technology, its industry and its target market.

Furthermore, this empirical research of the article does not incorporate the perspective of the managers of the business units that receive the innovation. Obtaining the vision from these profiles will enrich the data gathered regarding the research questions.

Further, the collected data were gathered from different countries. The corporate venture data were collected in Spain, Germany, and the UK, whereas spinoff data were collected in Finland. These countries have similarities and differences, and the latter might affect the findings of the article. Therefore, future studies could examine corporate venture and spin-off activities in the same country and the same industry.

Lastly, the findings of the article need to be confirmed through further research, something that creates room for future studies. For instance, it would be good to have a quantitative study where the identified activities are tested or a study of how the key activities impact the commercialization of technology-based companies' high-tech.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The corresponding author would like to thank the Design Processes Research Group from EINA University Center of Design and Arts (Autonomous University of Barcelona) for their support.

References

[1] C. Lettl, "User involvement competence for radical innovation", J. Eng. Technol. Manage., 24(1-2), 53-75, 2007. https://doi.org/10.1016/j.jengtecman.2007.01.004

[2] J. Nieto Cubero and C. Consolación Segura, "The Role Teams and Organizational Factors Play in Disruptive Innovations in Ambidextrous Organizations" in the 2019 ICE/ITMC International Conference on Engineering, Technology and Innovation, Valbonne Sophia-Antipolis, France, 2019, 1-14. https://doi.org/10.1109/ICE.2019.8792627

[3] S.A.Gbadegeshin, "The Commercialization Process of High Technologies: Case Studies from ICT, Cleantech and Life Sciences Industries" Ph.D. Thesis, University of Turku, Finland, 2019.

[4] X.M.Song and M.M.Montoya-Weiss, "Critical development activities for really new versus incremental products", J. Prod. Innovat. Manag., 15(2), 124-135, 1998. https://doi.org/10.1111/1540-5885.1520124

[5] C.M. McDermott and G.C. O'Connor, "Managing radical innovation: an overview of emergent strategy issues", J. Prod. Innovat. Manag., 19, 424–438, 2001. https://doi.org/10.1111/1540-5885.1960424

[6] S.A. Gbadegeshin, S.A, "Commercialization process of high technologies: case study of finnish university spin-off", Acad. Entrep. J., 23(2), 122, 2017.

[7] S.A. Gbadegeshin, "Stating best commercialization method: An unanswered question from scholars and practitioners", Int. J. Eco. Eng. Manag., 11(5), 1088-1094, 2017.

[8] M.A. Kirchberger, and L. Pohl, "Technology commercialization: a literature review of success factors and antecedents across different contexts", J. Technol. Transfer, 41(5), 1077-1112, 2016. https://doi.org/10.1007/s10961-016-9486-3

[9] K. Hindle and J. Yencken, "Public research commercialization, entrepreneurship and new technology-based firms: an integrated model", Technovation, 24(10), 793-803, 2004. https://doi.org/10.1016/S0166-4972(03)00023-3

[10] R.A.Burgelman and L.R. Sayles, "Inside corporate innovation: strategy, structure and managerial skills", Management, 18(2), 179-185, 1986. https://doi.org/10.3917/mana.182.0179

[11] S.A. Hill and J. Birkinshaw, "Strategy-organization configurations in corporate venture units: Impact on performance and survival", J. Bus. Venturing, 23(4), 423-444, 2008. https://doi.org/10.1016/j.jbusvent.2007.04.001

[12] E. van Burg, S. de Jager, I.M.M.J. Reymen and M.Cloodt, " Design principles for corporate venture transition processes in established technology firms", R&D Manage., 42(5), 455-472, 2012. https://doi.org/10.1111/j.1467-9310.2012.00695.x

[13] M.P. Rice, R. Leifer, and G.C. O'Connor, "Commercializing discontinuous innovations: bridging the gap from discontinuous innovation project to operations", IEEE T. Eng. Manage., 49(4), 330-340, 2002. doi: 10.1109/TEM.2002.806721

[14] R. Garcia and R. Calantone, "A critical look at technological innovation typology and innovativeness terminology: a literature review", J. Prod. Innovat. Manag., 19(2), 110–132, 2002. https://doi.org/10.1111/1540-5885.1920110

[15] V. Govindarajan and P.K. Kopalle, "The usefulness of measuring disruptiveness of innovations ex-post in making ex-ante predictions", J. Prod. Innovat. Manag., 23, 12–18, 2006. https://doi.org/10.1111/j.1540-5885.2005.00176.x

[16] D. Dougherty, "Understanding new markets for new products", Strateg. Manage. J., 11, 59-78, 1990. https://doi.org/ 10.1108/00251749810223592

[17] R.A. Burgelman, "A process model of internal corporate venturing in the diversified major firm", Admin. Sci. Quart., 28(2), 223-244, 1983. https://doi.org/10.2307/2392619

[18] S. Ford, E. Garnsey and D. Probert, "Evolving corporate entrepreneurship strategy: technology incubation at Philips", R&D Manage., 40(1), 81-90, 2010. https://doi.org/10.1111/j.1467-9310.2009.00580.x

[19] G.C. O'Connor, and R. de Martino, "Organizing for radical innovation: An exploratory study of the structural aspects of RI management systems of large established firms", J. Prod. Innovat. Manag., 23, 475-497, 2006. https://doi.org/10.1111/j.1540-5885.2006. 00219.x

[20] M.L. Tushman and C.A. O'Reilly, "Ambidextrous organizations: managing evolutionary and revolutionary change", Calif. Manage. Rev., 38(4), 8-30, 1996. https://doi.org/10.2307/41165852

[21] D.A. Levinthal and J.G. March, "The myopia of learning", Strateg. Manage. J., 14, 95-112, 1993. https://doi.org/10.1002/smj.4250141009

[22] H. Chesbrough, and C.L. Tucci, "Corporate Venture Capital in the Context of Corporate Innovation". EPFL Working Paper, Lausanne, 2005.

[23] R.G. McGrath, "Exploratory learning, innovative capacity, and managerial oversight", Acad. Manag. J., 44(1), 118-131, 2001. https://doi.org/10.5465/3069340
[24] N. Thornberry, "Corporate entrepreneurship: Antidote or oxymoron?", Eur. Manag. J., 19(5), 526-533, 2001. https://doi.org/10.1016/S0263-2373(01)00066-4
[25] E. G. Carayannis, E.M. Rogers, K. Kurihara and M.M. Allbritton, "High-technology spin-offs from government, R&D laboratories and research universities", Technovation, 18(1), 1-11, 1998. https://doi.org/10.1016/S0166-4972(97)00101-6
[26] F. Pirnay, B. Surlemont and F. Nlemvo, "Toward a typology of university spin-offs", Small Bus. Eco., 21(4), 355-369, 2003. https://doi.org/10.1023/A:1026167105153
[27] S. Shane, Academic Entrepreneurship: University Spin-offs and Wealth Creation., UK, Edward Elgar Publishing Ltd, 2004.
[28] G.S. Lynn and A.E. Akgün, "Innovation strategies under uncertainty: A contingency approach for new product development", Eng. Manag. J., 10(3), 11-18, 1998. https://doi.org/10.1080/10429247.1998.11414991
[29] O. Gassmann, B. Widenmayer and M. Zeschky, "Implementing radical innovation in the business: the role of transition modes in large firms", R&D Manage., 42(2), 120-13, 2012. https://doi.org/10.1111/j.1467-9310.2011.00670.x
[30] R. Chandy and G. Tellis, "The incumbent's curse? incumbency, size, and radical product innovation", J. Marketing, 64(3), 2000. https://doi.org/ 10.1509/jmkg.64.3.1.18033
[31] G.C. O'Connor, R. Hendricks and M.P. Rice, "Assessing transition readiness for radical innovation", Res. Tech. Manage., 45(6), 50-56, 2002. https://doi.org/10.1080/08956308.2002.11671533
[32] D. Clausing and M. Holmes, "Technology readiness", Res. Tech. Manag., 53(4), 52-59, 2010. https://doi.org/10.1080/08956308.2010.11657640
[33] J.C. Mankins, "Technology readiness and risk assessments: A new approach", Acta Astronaut., 65(9), 1208-1215, 2009. https://doi.org/10.1016/j.actaastro.2009.03.059
[34] C.A. O'Reilly and M.L. Tushman "The ambidextrous organization", Harvard Bus. Rev., 82(4), 74-81, 2004.
[35] G.C. Gilbert, "Change in the presence of residual fit: Can competing frames coexist?", Organ. Sci., 17(1), 150–167, 2006. https://doi.org/10.1287/orsc.1050.0160
[36] J.J.P. Jansen, M.P. Tempelaar, F.A.J. van den Bosch and H.W. Volberda, "Structural differentiation and ambidexterity: The mediating role of integration mechanisms", Organ. Sci., 20(4), 797-811, 2009. https://doi.org/10.1287/orsc.1080.0415
[37] C. Andriopoulos and M.W. Lewis, "Exploitation-exploration tensions and organizational ambidexterity: Managing paradoxes of innovation", Organ. Sci., Special Issue on Ambidextrous Organizations, 20(4), 696-717, 2009. https://doi.org/10.1287/orsc.1080.0406
[38] O.P. Kauppila, "Creating ambidexterity by integrating and balancing structurally separate inter-organizational partnerships", Strateg. Organ., 8(4), 283-312, 2010. https://doi.org/10.1177%2F1476127010387409
[39] S. Raisch and M.L. Tushman, "Growing new corporate businesses: From initiation to graduation", Organ. Sci., 27(5), 1237-57, 2016. https://doi.org/10.1287/orsc.2016.1081
[40] D.A. Garvin and L.C. Levesque, "Meeting the challenge of corporate entrepreneurship", Harvard Bus. Rev., 84(10), 102-112, 2006.
[41] M.L. Tushman and D. Lavie, "Exploration and exploitation within and across organizations", Acad. Manag. Ann., 4(1), 109-155, 2010. https://doi.org/10.5465/19416521003691287
[42] M. Steffensen, E.M. Rogers and K. Speakman, "Spin-offs from research centers at a research university", J. Bus. Venturing, 15(1), 93-111, 1998. https://doi.org/10.1016/s0883-9026(98)00006-8
[43] F. Frattini, A. De Massis, V. Chiesa, L. Cassia and G. Campopiano, "Bringing to market technological innovation: What distinguishes success from failure", Int. J. Eng. Bus. Manag., 4(15), 1 -11, 2012. https://doi.org/10.1111/j.1540-5885.2006.00220.x
[44] T. M. Nevens, G. L. Summe, and B. Uttal, "Commercializing technology: What the best companies do", Harvard Bus. Rev., 64(8), 154–163, 1990. https://doi.org/10.1108/eb054310
[45] J.S. Gans and S. Stern, "The product market and the market for "ideas": commercialization strategies for technology entrepreneurs", Res. Policy, 32(2), 333-350, 2003. https://dx.doi.org/10.2139/ssrn.317219
[46] J. Pellikka and P. Malinen, "Developing Commercialisation of Innovation in High Technology Industries – Regional Perspective", in 2011 56th International Council for Small Business (ICSB), Stockholm, Sweden, June 15–18, 2011.
[47] P. Valiauga, "Commercialization of High-tech Radical Innovations: Case Studies of X-ray Imaging Technologies." in 2013 paper presented at the Aalto University School of Science, Finland, May 16, 2013. http://noppa.aalto.fi/noppa/kurssi/tu-22.1500/luennot/TU-22_1500_povilas
[48] S.P. Nichols, "Module 1: An Introduction to Commercialization of Science and Technology. Converting Technology to Wealth Workshop". Austin, TX: IC2 Institute, The University of Texas at Austin. http://ut.gtrade.or.kr/inc/download.asp?key=5288
[49] A. Dhebar, "Bringing new high-technology products to market: Six perils awaiting marketers", Bus. Horiz., 59, 713—722, 2016. https://doi.org/10.1016/j.bushor.2016.08.006
[50] A.D. Chandler, "Commercializing high-technology industries", Bus. Hist. Rev., 79(3), 595–604, 2005. https://doi.org/ 10.1017/S0007680500081460
[51] A.C. Fletcher and P.E. Bourne, "Ten simple rules to commercialize scientific research", PLoS Comput. Biol., 8(9), e1002712, 2012. https://doi.org/10.1371/journal.pcbi.1002712
[52] S. F. Slater and J. J. Mohr, "Successful development and commercialization of technological innovation: insights based on strategy type", J. Prod. Innovat. Manag., 23(1), 26–33, 2006. https://doi.org/10.1111/j.1540-5885.2005.00178.x
[53] E. Maine and E. Garnsey, "The commercialization environment of advanced materials ventures", Int. J. Technol. Manage., 39(1/2), 49 -71, 2012. https://doi.org/ 10.1504/IJTM.2007.013440
[54] J. E. Amadi-Echendu and R. T. Rasetlola, "Technology Commercialization Factors, Frameworks and Models" in 2011 IEEE/ITMC Conference, 144-148, 2011.
[55] D. S. Rogers, D. M. Lambert, and A. M. Knemeyer, "The Product Development and Commercialisation Process", Int. J. Logist. Manag., 15(1), 43 – 56, 2004. https://doi.org/10.1108/09574090410700220
[56] J. B. Pietzsch, L. A. Shluzas, M. E. Paté-Cornell, P. G. Yock and J. H. Linehan, "Stage-gate process for the development of medical devices", J. Med. Devices, 3(021004), 1-15, 2009. https://doi.org/10.1016/j.technovation.2017.03.003
[57] T. Elfring and N. Foss, "Corporate Renewal Through Internal Venturing and Spinn-offs: Perspectives from Organizational Economics", No 97-7, IVS/CBS Working Papers, 97-7, 1997.
[58] B. L. Golish, M.E. Besterfield-Sacre and L. J. Shuman, "Comparing Academic and Corporate Technology Development Processes", J. Prod. Innovat. Manag., 25(1), 47-62, 2008. https://doi.org/10.1111/j.1540-5885.2007.00282.x
[59] P.P. Georgallis and R. Durand, "Achieving High Growth in Policy Dependent Industries: Differences between Startups and Corporate-Backed Ventures", Long Range Plann., 50(4), 487-500, 2017. https://doi.org/10.1016/j.lrp.2016.06.005
[60] K. Wennberg, J. Wiklund and M. Wright, "The Effectiveness of University Knowledge Spillovers: Performance Differences between University Spinoffs and Corporate Spinoffs", Res. Policy, 40(8), 1128-1143, 2011. https://doi.org/10.1016/j.respol.2011.05.014
[61] S.A. Zahra, E. Van de Velde and B. Larraneta, "Knowledge Conversion Capability and the Growth of Corporate and University Spinoffs", Ind. Corp. Change, 16(4), 569–608, 2007. https://doi.org/10.1093/icc/dtm018
[62] R. C. Shrader and M. Simon, "Corporate Versus Independent new Ventures: Resource, Strategy, and Performance Differences", J. Bus. Venturing, 12(1), 47-66, 1997. https://doi.org/10.1016/S0883-9026(96)00053-5
[63] T. Bager, H. Ottosson and T. Schøtt, "Intrapreneurs, Entrepreneurs and Spin-off Entrepreneurs: Similarities and differences", Int. Small. Bus. J., 10(3), 2010. https://doi.org 10.1504/IJESB.2010.033572
[64] R.K. Yin, Case Study Research, Thousand Oaks, CA: Sage Publications, 1994.
[65] E. Danneels, "Disruptive technology reconsidered: a critique and research agenda", J. Prod. Innovat. Manag., 21(4), 246–258, 2004. https://doi.org/10.1111/j.0737-6782.2004.00076.x
[66] K.M. Eisenhardt, "Building theories from case study research", Acad. Manag. Rev., 14(4), 532-550, 1989. https://doi.org/ 10.2307/258557
[67] D. E. Gray, Doing Research in the Real World, Chapter 2: Theoretical Perspectives and Research Methodologies, USA: Sage Publications, 2013.
[68] J. W. Creswell, Research Design: Qualitative, Quantitative and Mixed Methods Approaches. 3rd Ed., United Kingdom, London: Sage Publication, 2009.
[69] P. Eriksson and A. Kovalainen, Qualitative Methods in Business Research, 1st Ed., United Kingdom, London: Sage Publications, 2008.
[70] S. Rajasekar, P. Philominathan and V. Chinnathambi, V.Rese. Method., 2016. http://arxiv.org/pdf/physics/0601009.pdf
[71] N. Denzin and Y. Lincoln, Handbook of Qualitative Research, United Kingdom, London: Sage Publications, 2000.
[72] J.M. Corbin and A. Strauss, Basics of Qualitative Research, 3rd Ed., Thousands Oaks: Sage Publications, 2008.

[73] J. Nieto, “Managing Integration of Disruptive Innovations in Ambidextrous Organizations”, in 2018 ISPIM Innovation Conference, Stockholm, Innovation, the Name of the Game, Stockholm: LUT Scientific and Expertise Publications, 2018.
[74] N. Abd Rahima, Z. B. Mohamed, and A. Amrin, “Commercialization of emerging technology: The role of academic entrepreneur”, Soc. Behav. Sci., 169, 53 – 60, 2005. https://doi.org/10.1016/j.sbspro.2015.01.285
[75] E. W. Eldred and M. E. McGrath, “Commercializing new technology”, Res. Tech. Manag., 40(1), 41-47, 1997. https://doi.org/10.1080/08956308.1997.11671102
[76] J. E. Amadi-Echendu and A. John, “On the Commercialization of Trailing Knowledge and IP”, in 2008 PICMET’08 Conference, Cape Town. Paper ID: 08R0018, 2008.
[77] A. Al Natsheh, S.A. Gbadegeshin, A. Rimpiläinen and T. Mainela, "Technology-based entrepreneurship: measurement technology perspective", Int. J. Res. Bus., 2(9), 26 – 42, 2013.
[78] J. Che and R. Panda, “Review: Commercialisation of Piezoelectric Single Crystals for Medical Imaging Applications”, in 2005 IEEE Ultrasonics Symposium, 235-240, 2005.
[79] H. Chesbrough, “Business model innovation: opportunities and barriers”, Long Range Plann., 43(2), 354-363, 2010. https://doi.org/10.1016/j.lrp.2009.07.010
[80] D.J. Teece, “Business models, business strategy and innovation”, Long Range Plann., 43(2), 172-194, 2010. https://doi.org/10.1016/j.lrp.2009.07.003
[81] C. Zott, R. Amit and L. Massa, ”The business model: recent developments and future research”, J. Manage., 37(4), 1019-1042, 2011. https://doi.org/10.1177%2F0149206311406265

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 635-645 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Activerest: Design of A Graphical Interface for the Remote use of Continuous and Holistic Care Providers

Nuno Martins[1,*], Jéssica Campos[2], Ricardo Simoes[2]

[1]*Polytechnic Institute of Cavado and Ave / ID +, Barcelos, Portugal*

[2]*Polytechnic Institute of Cavado and Ave, Barcelos, Portugal*

ARTICLE INFO

Article history:
Received: 14 January, 2020
Accepted: 11 March, 2020
Online: 14 April, 2020

Keywords:
Pressure ulcers
Mobile application
Digital technologies
Healthcare mobile solution
Android

ABSTRACT

One of the most prevalent physical limitations of both aging and accidents is immobility. Being bedridden requires great care to avoid the formation of pressure ulcers. Thus, the 'ActiveRest' project aimed to develop a textile mattress guard that will contribute to the prevention of pressure ulcers. This textile mattress guard integrates a monitoring system, based on a network of sensors (pressure, temperature, humidity), coupled with an intelligent non-invasive actuation system. The major innovation is that the actuation system only acts around high pressure loci, effectively relieving pressure in those critical areas. This system can operate autonomously and continuously adjusts its response to the data obtained by the real-time pressure sensor network. In addition, the system encompasses a mobile application that communicates in real-time with the mattress guard, and that enables caregivers to monitor the condition of a patient (and even receive an alert whenever their direct action is required). In the scope of ActiveRest, this paper describes the work pertaining to the digital technology component of the system, namely the graphic design and the development of the App necessary to convey the data to the caregiver. In this case, the main focus was the design of a graphic user interface, simple and customized towards distinct target audiences, which will facilitate the caregiver's work and interaction with the pressure ulcer prevention system.

1. Introduction

In recent decades there has been a growth in average life expectancy in developed countries, according to a study published in The Lancet magazine, in 2020 for the first time in history there will be more people aged 65 years or more than aged 5 years or less [1].

In Portugal, in 1970 a person aged 65 could live on average more 13.5 years. In 2017, this number has risen to 19.5 years [2]. However, in 2013, the years considered healthy after 65 years of age had an average of 9.2 years.

Despite the increase in average life expectancy, aging still causes several types of limitations to the human body, it begins at 25 years and after the age of 60 accelerates its process which causes more physical and cognitive difficulties to the individual. In addition to aging, emerging diseases are also a factor that increases these limitations [3].

Immobility is one of the examples of physical limitations caused by aging and more prone diseases among the elderly. In addition to numerous complications, one of the factors aggravated by it is the emergence of pressure ulcers (PU) which, according to the European Pressure Ulcer Advisory Panel (EPUAP), are defined as "a localized lesion of the underlying skin or tissue, usually over a bony prominence as a result of pressure or a combination of bone and torsional forces. " [4]. In most cases, the emergence occurs when the patient is bedridden and has little or no mobility, which requires him or her to be accompanied by a health professional or family member to move to prevent PU.

Faced with this problem and given the seriousness of the emergence of a pressure ulcer for the patient and caregiver, as well as the economic burden on the national health services (NHS), there is a need to implement preventive measures. as well as timely therapeutic means to face the risk of pressure ulcers development.

So, the ActiveRest project, to which this paper is part, proposes a solution that can act on time to prevent this type of skin injury, while providing immediate relief from pressure on the user. This solution is composed of two components that complement each other, although they can act independently. One of these components is a textile guard that incorporates a sensing system composed of pressure, temperature and humidity sensors, three

*Corresponding Author: Nuno Martins, nunomartins.com@gmail.com

risk factors for the development of PUs, together with an autonomous localized pressure relief system, which is based on an actuation mechanism. noninvasive, light and compact. The other component, which represents one of the innovative features of the proposed solution and which makes it stand out from the existing solutions in the market, constituting the central focus of this work, is the design of a digital graphic interface through which the user can interact with the product. safeguard, namely, access and manage the data derived from its performance.

Thus, this work aims to develop a Mobile Application (app), which will communicate with the guard via wireless, and which will enable the caregiver to check the patient's condition at any time. In this case, it is intended to implement two versions of the app that are suitable for patient monitoring in more technical or familiar environments (hospital / institutional or home environment). Thus, a home version was designed for use by family members and caregivers in the home environment. And another version designed for the institutional / hospital environment for use by doctors, nurses and other professionals/caregivers who have direct contact with patients, this version may be used in institutions such as nursing homes, retirement homes, rehabilitation centers / continuous care/palliative care, as well as in hospitals and other units that receive this category of patients.

Figure 1 presents an infographic perspective of the underlying approach taken by the ActiveRest project.



Figure 1: ACTIVEREST Project

The solution presented in this project, in addition to designing an application that is capable of displaying data from communication with the bedridden, also has a set of other features that complement the patient's treatment and give the application another level of protection. So, based on what has been said, the research carried out and presented in this article aimed to improve the development of a mobile application in order to design a more versatile and efficient solution that enables better results of the mobile application in the use of textile protection by the digital component.

2. Digital Transformation of Health in Portugal

Digital transformation is increasingly present in health and, in this sense, "several systems have been developed, implemented and updated to improve the quality of the information in healthcare provision and management, increase efficiency and technologically upgrade existing systems. , allowing for an improvement in public services provided to citizens [...]" [5].

In Portugal, the eHealth Summit event has been held in recent years where the main advances in the Health area are debated. Besides, several implementations of the concept have been observed in the country, an example is the Paperless Receipt (PR) which, according to the NHS, consists of the following: "With PR, the user may provide his doctor with an email and mobile number to receive prescribing information. At the time of consultation, he will be able to receive the treatment guide in paper form, or choose to receive it via email, and also access and consult it in the Citizen Area" [6]. In August 2018, the PR represented 97.02% compared to the traditional ones [7]. This figure reflects the effectiveness of the implementation of the system which, besides reducing costs and contributing to environmental preservation, has greater safety and control of the prescription of the medicine.

Henrique Martins, President of the Ministry of Health's Shared Services (HSS), states that "[...] we cannot continue to program and create digital solutions, assuming that people are sitting in front of a fixed terminal. People move. Healthy people do not go to the NHS much, nor do they want or seek health care… It is health that has to go after them and preferably in their pockets and everywhere. Therefore, all mobile digital transformation requires a paradigm shift in programming, creation, and the relationship of digital tools." [8, p. 9]. In Portugal was created MySNS Carteira, an app that allows the user to consult his health information present in the National Register of User (NRU), through his own mobile phone. In May 2018 this was the most popular app in Portugal in the Health & Fitness category of the GooglePlay and AppleStore stores [9, 10].

MySNS Carteira is an example of mHealth - Mobile Health that "refers to the practice of medical and public health supported by mobile devices such as mobile phones, tablets, personal digital assistants (PDA) and wireless infrastructure. [...] mHealth covers all applications of telecommunications and multimedia technologies for providing health information." [11].

In the Portuguese context, mobile applications are the most used medium in mHealth, with surveillance, monitoring and information exchange being the most used features in the development of these apps [12].

Through several initiatives, such as those mentioned above, Portugal distinguishes itself at the forefront of Digital Health Transformation with prominence among the member countries of the European Union [5]. Despite this achievement, SPMS President Henrique Martins has listed eight essential conditions for the country to continue with this advance: (1) Social and political motivation and support; (2) eSkills of health professionals and digital literacy of citizens; (3) Technical skills of technical staff in the State and in companies; (4) Strengthening the action of digital transformation agents; (5) Robust infrastructure, fast networks, and appropriate hardware; (6) Digital pillars and experience; (7) Mobile-to-start principle; (8) All2All thinking; CyberSecurity as a Service; TeleHealth as a principle rather than as a supplement [8].

Under these conditions, it is possible to observe that both health professionals and other citizens are fundamental elements for TDS in Portugal to continue to happen gradually. Through the support, receptivity, and use of technological systems, it will be possible to create a means of collaboration for this advance. And for these technologies to be successful before the public, the intervention of professionals specialized in technological and digital areas, such as programmers and designers, is required.

3. Mobile Health Applications

Since its inception, the use of mobile applications has spread. The massification of mobile devices coupled with the expansion of easy access to the internet, justify the growing increase in their use. Other factors that contributed to this growth are the ability to be developed for various types of needs, and the mobility characteristic of mobile devices. The number of app store downloads is growing rapidly every year. In 2017 the number of downloads was 178.1 billion, the forecast is that this number will increase to 258.2 billion in 2022 [13].

Wireless technologies combined with mHealth enable healthcare to be increasingly connected. In the digital health market, mHealth has been the fastest growing technology since 2015 [14]. The diversity of mHealths present in online stores is extensive, being one of the fastest-growing categories. According to a statistical analysis by Statista, the Health & Fitness group is AppStore's eighth most popular group.

Through the use of wearable equipment (electronic devices that capture and send information and can be used as clothing accessories, such as a bracelet or watch) apps are developed for various uses and mainly for monitoring, storage and data exchange. FitBit is a company that manufactures a wearable technology-based wristband that collects fitness data from the wearer and uploads it to their smartphone storage. According to one study, the use of the bracelet proves to be effective in weight loss [15]. In addition, other research indicates that data collection by the device is accurate and reliable [16]. In 2016, the same company was in the top 50 on the list of the 500 fastest growing companies worldwide [17]. Just like the FitBit bracelet, there are many other solutions in this context, which have a good reception by the public.

Mobile applications have changed the global health scenario in both work and personal care. There is a wide variety of applications focused on this area, and the trend is its gradual increase, especially with the introduction of new devices that, together with smartphones, allow the development of increasingly accurate and new mHealths. functionalities. On the other hand, because it involves a delicate subject such as health, it is necessary that these applications have supervision by a regulatory body, especially those that indicate diagnosis, since the apps are not 100% safe, always requiring the review of a professional. Cybersecurity is also an important issue, namely in protecting the user's personal data.

Based on the objectives of the ActiveRest project, a mobile application was developed which, through data collection and exchange with the textile guard, enables better patient monitoring and, consequently, the prevention of pressure ulcers. Although there are already apps on the market that focus on the same theme of this research, they do not meet the needs of the equipment in question, for two main reasons: first because it is a device that has been developed since its inception, featuring certain particularities and functionalities that set it apart from other equipment; and second for distancing itself from apps with unattractive aesthetics to the user, being one of the main goals the development of an app with an appealing aesthetics for different segments of the app users.

4. Research Contribution

Pressure ulcers are classified by the NHS in the National Network of Integrated Continued Care group (RNCCI) which “has as its main objectives the provision of continuous and integrated health care and social support to people who, regardless of age, are in a situation of dependence and loss of autonomy” [18, p. 109]. Detailed information on health statistics and the burden of pressure ulcers on the economy has been provided by OCDE [19].

The DGS indicates that about 95% of pressure ulcers are preventable, as long as their degree of risk is identified in advance, so one of the factors that help prevent this problem is proper patient monitoring. One of the many problems that arise from the emergence of pressure ulcers is the increased length of stay of the patient causing the NHS greater financial burden.

The development of a mobile application that allows the caregiver to check the patient's condition at any time could improve the prevention of pressure ulcers in order to optimize patient monitoring through the guarding sensing system which will send data about your performance for the app. With the ease of mobility of mobile devices, healthcare professionals, through the use of guard and mobile applications, will have greater control over the risk of developing pressure ulcers in patients using this type of equipment.

Through the use of the app, the workload of professionals responsible for providing the necessary care to these patients tends to decrease, since the system works autonomously, alerting the caregiver whenever there is a need for repositioning. as well as which or which body regions are under pressure. The decrease in the appearance of pressure ulcers, in addition to bringing more quality of life to the patient and their caregiver, will also present to the NHS a decrease in the expenses resulting from this health problem.

5. State of Art

For this analysis, we used mobile applications and device interfaces that integrate objects that act physically in the prevention/treatment of ulcers, such as mattresses with pressure sensors. The examples below correspond to existing applications/interface models that are designed to assist patients and their caregivers in the prevention and treatment of pressure ulcers, and also serve as an informative tool for this type of skin wound.

Applications and software analyzed: A) VA Pressure Ulcer Resource mobile application, B) XSENSOR - FORESITE PT software, C) uControl app and D) Lifelens.

The VA Pressure Ulcer Resource application developed for pressure ulcer and injury prevention is intended for US war veterans and their caregivers. Features such as history; sticky notes; notes; pressure and time monitoring graphs; alerts for the user to move around; a shortcut to get urgent care, as well as knowing where the nearest hospital is. These features allow the user to act proactively, through history and reminders, preventing the appearance of pressure ulcers. Alerts allow mobile users or

caregivers to act at certain times, helping the patient to move so that there are no very long periods of body immobility [20].

The XSENSOR – FORESITE PT created ForeSite PT, which is a body pressure monitoring system using a mattress cover and a touch screen monitor with software that analyzes and stores data. Real-time display of a body map of the pressure being exerted allows the user to act on body movement only when necessary. Graphs with the duration of the current position and the occurrence of repositioning allow the system to be used by different users, a common occurrence in hospitals. The timer indicates the time intervals to be reset and is positioned prominently on the interface because it is a key feature of the system, allowing the user greater control over patient movements [21].

Unitron is a multinational company with high experience in hearing solutions. As a member of the Sonova Group, Unitron develops high-performance hearing solutions that enable natural listening to sounds and speech. In addition to hearing aids, Unitron offers its users three applications: Remote Plus, uControl and Start Here. These applications allow the user to get the most out of the device by adjusting the volume and balance of the hearing aids. Constant assessment of user satisfaction is also allowed, as is communication between users. This application is developed for the senior public and differs from the vast majority found in this market by presenting an attractive and intuitive design, managing through graphic minimalism to effectively transmit messages to the UI [22].

Pressure Ulcer Prevention is a mobile application developed by the University of Southern California School of Medicine for patients with pressure ulcers to use as a follow-up, information and research about the disease, thus allowing the sharing of information with doctors or caregivers for better identification of causes and subsequent prevention. The application allows the user to photograph and annotate some parts of his body - specifically the areas most sensitive to the appearance of pressure ulcers - for monitoring. With this data, the app creates graphs showing the history of the monitored part. It also allows the user to create alerts to remind them to reposition themselves to avoid constant pressure on a certain area of the body. It has a didactic component where it tells the patient recommendations to reduce the risk of ulcers. And just like the previous mobile app, it even includes a shortcut for the user to make emergency calls [23].

Vista Medical is a company that develops technologies in the field of sensors and pressure mapping. This brand has developed the BodiTrack system that consists of a kind of sensor blanket and a USB cable for the integration of devices such as tablets and computers to access the product software.

This software allows the user to view the history of pressure exerted at a given time; visualization of areas with higher pressure; indicates the risk factor of each body part according to sensing; allows you to time the repositioning time; indicates the maximum pressure exerted during a selected period; ability to create alarms and real-time pressure mapping [24].

Lifelens is a Windows Phone 7 application that diagnoses malaria with 94% accuracy, which contrasts with 40% of conventional methods [25]. This app significantly reduces the costs of testing and traveling teams for testing, as it only requires the user to photograph blood already combined with special malaria dye, then a specific algorithm indicates if the patient is infected. Results can be sent via the Internet, along with GPS coordinates, thus enabling faster interventions in disease outbreaks.

According to this analysis, it was possible to identify some functions that were the basis for the ActiveRest app project, namely: the real-time visualization of the pressure exerted by the body. This feature quickly and effectively provides you with information relevant to pressure ulcer prevention along with history, reminders, notes, pressure and time monitoring graphs.

The most relevant features for the study of the creation of the ActiveRest app were: the real-time visualization of the pressure exerted by the body to the local and general software. pressure ulcer prevention; pressure and repositioning history for better analysis of patient behavior; pressure and time monitoring reminders, diaries, and graphs for optimizing information catalysis.

An important detail found in the analyzed applications (table 1) was the presence of an emergency button that allows the user to quickly contact another person. However, it was observed that despite having an interesting concept, shortcuts in applications are confusing and with various actions to make the contact, causing the functionality to lose its main purpose. In ActiveRest we analyzed a more effective way to make this contact, allowing the user to perform the task with the fewest possible actions.

In the analysis of solutions for the prevention and treatment of pressure ulcers, it can be observed that the interfaces need a specific development, as well as the attention regarding UX because it is a product to be used by two audiences of different age groups. One of the challenges of this project was the development of interfaces that meet the needs of both segments.

Table 1: Comparative table of applications and software functionalities.

	APP - SOFTWARE				
FUNCTIONALITY	VA PRESSURE ULCER RESOURCE	PRESSURE ULCER PREVENTION	XSENSOR FORESITE PT	VISTA MEDICAL BODITRACK	UNITRON
Body Pressure Map					
Body Humidity Map					
Historic					
Sticky notes					
Daily					
Monitoring					
Data Graph					
Move Alert					
Emergency button					
UP's Information					
Stopwatch					

PRESENT

ABSENT

6. Mobile Application Development

Extensive analytical work on the literature on the types of mobile applications and graphic software interfaces in the area of pressure ulcer prevention devices / systems has culminated in an app proposal and its specifications have been defined. In this way, it was possible to define the necessary steps for the development of the app, among which the characterization of the people; the information architecture; wireframes and wireflows; the visual identity; prototyping; and the mobile application to the current development of textile protection.

6.1. Personas

Persona 1 - Bruno, 30, male nurse responsible for providing care and monitoring bedridden patients. He works in a large hospital and looks after a sector with 25 bedridden patients, some of whom have pressure sores and the remaining developmental risks. It uses the ActiveRest app along with the textile guard to monitor patients more effectively, enabling them to make better and more accurate patient surveillance and, when necessary, to take appropriate preventive measures early. It also uses the app to record and manage patient information, including repositioning history, humidity, pressure, temperature, and view notes and reminders.

Persona 2 - Joana, 42, journalist, responsible for accompanying her bedridden mother at home. Joana works at home but sometimes needs to be absent. Use the app and guard to know when to reposition her mother as well as to receive alerts when an abnormal or hazardous situation is detected, including critical pressure spikes and/or excessive temperature and humidity levels.

Persona 3 - Maria, 67, retired, has reduced mobility, so she spends many hours in bed. In addition to lack of mobility and sensitivity, which requires help to move and change her position, she also has memory problems, having some difficulty remembering to take the medicines. That said, Maria uses the app as a resource in her care, taking advantage of the app's alarmistic ability to reposition herself, as well as medication reminders, among other features that help her in her daily life and that provide her with preventive measures against her condition. She also uses the diary feature and reminder setting to create notes that may be important to present to the attending physician. She also makes use of the emergency call function as she has an imminent risk of myocardial infarction.

6.2. UX UI for audiences of different age

A aplicação móvel será utilizada por públicos de diferentes idades, sendo a idade inicial os 25 anos. O público jovem/adulto geralmente possui melhor literacia digital em comparação ao público idoso (65+). Por este motivo, foi necessária uma investigação a respeito do desenvolvimento de interfaces digitais para públicos com diferentes níveis de literacia digital.

The worldwide increase in the number of older people in an increasingly digital society requires that digital technology solutions were increasingly adapted to this public. The aging of the human body, even if currently possible, is one of the main obstacles in the use of digital media, due to the gradual loss of motor and mental abilities. However, the lack of use of appropriate solutions for this class of users in the development of digital interfaces is also one of the major factors for the frustration and withdrawal of the use of these technologies.

Data indicate that the usability of websites decreases 0.8% per year between the ages of 25 to 60 years, the reason is the increased time to complete each task on the page, and difficulties in interacting with the interface. This percentage tends to increase after age 60, a 75-year-old is 74% slower to use websites than a conventional 35-year-old. This is due to the consequences of human aging that contribute to it, among them it is possible to mention: the decrease of cognitive resources; of visual ability; the ability to memorize; of perception; and motor skills [3].

Nielsen says that designing only for the young audience may result in the elderly giving up the interface. In contrast, when designing for the elderly the chances of the interface being used by both audiences are significantly higher because the basis of the research will be for the development of an easily accessible interface for people with reduced mental and motor skills, so the interface can answer also to the young public that usually has these capacities in better condition, However, good practices for aesthetic design must be present in the project, as it is an important the attractive component for young audiences. Thus, the adaptability of the mobile application is an important resource to be used for both audiences to achieve greater success in using the app.

According to an analysis of several studies, it was possible to find different recommendations to be followed so that the interaction of senior users with digital interfaces is of the least possible frustration. These recommendations are based on human behaviour in a digital environment as they age, and the issues of decreased vision, motor coordination, and reduced memory capacity are the main factors for their development.

These recommendations, although based on the development of interfaces for the elderly, are largely similar to those indicated in the design for the adult audience. About the 10 Interface Design Usability Heuristics developed by Nielsen, the author states that: “The 10 Principles of Interface Design Usability are called heuristics because they are general rules rather than specific usability guidelines” [4].

With this research, it was possible to compile the analysed studies on the recommendations for interface design for seniors and then created a table where it is possible to briefly find these guidelines (Table 2). This table aims to facilitate and speed up the design of interfaces for seniors, so that following the suggestions in it is likely that the final product has a better solution for the user.

These recommendations have been used in the implementation of both versions of the app, so that both classes of users can use both versions of the app without causing them frustration or major handling difficulties.

6.3. Information Architecture

Information Architecture (AI) is a term created by designer Richard Saul Wurman. In his book “Information Architects” he states that the information architect is the “individual who organizes patterns inherent in data, transforming the complex into something clear” and who “creates the structure or map of certain information to enable others to create their path toward knowledge.” It further states that “AI is the emerging profession of the 21st century, whose scope is shaped by current needs, focused on clarity, human understanding and the science of information organization.” [26]

Table 2: Recommendations for the design of digital interfaces for the elder

NAVIGATION	TEXT	COLOR	GRAPHISM
Home page customization option;	Font size should not be less than 14pt;	Color contrast: 4.5: 1 or 3: 1, depending on font size;	Visual noise reduction;
Allow the user to enlarge the screen;	Use non-serif and non-condensed typography whenever possible;	Option to change the colors of the application, and the other version of the theme should have colors with more contrast;	The application icon must be easily recognizable;
Always inform the user of the current page;	80 characters per line in text;		Try to use familiar and user-recognizable symbols;
Minimum possible actions to perform tasks;	Avoid the use of justified text;	Actions when identified by colors should also, whenever possible, make use of text or symbols;	Buttons with clear language and distinct from the other features of the app and avoid having too small dimensions, with the recommended minimum size 9mm high or wide;
Avoid drastic changes to the application interface when updates occur;	Use few font styles as well as different types;	Avoid using colors such as blue and green on the same screen, as the elderly have difficulty distinguishing them in screens;	
Whenever there is an error-prone and irreversible action, ask the user for task verification and confirmation;	Create option to increase or decrease mobile application font size;		
The most important features should be visible on the first page avoiding the need for many interactions to find them;	Make use of clear and simple language that does not force the user to make too much memorial effort to identify an action;	In the background of texts avoid high contrast and opt for a pastel color so that there is no interference in reading due to the screen light;	
Make navigation as predictable as possible, as well as consistent with using the same features across all app pages;	Avoid the use of unnecessary words, where possible use a word to describe a function, always simply and clearly;		
Whenever possible, insert top and bottom navigation bars should be uniform throughout the app navigation and should have frequently used page access buttons such as Home;	Make clear the hierarchy of information, clearly and clearly exposing the titles, and the information, according to its degree of importance, should be highlighted or evidenced;		

In the case of a mobile application, the information architecture brings together the concepts that will be present in the project and logically segments the information that will give rise to the features to be inserted in the app. This step makes the development of functions more objective and understandable. It is possible to observe in Figure 2, the information architecture for the app that is intended to be developed within the ActiveRest project, through a demo with the main specifications of the app, taking into account the different scenarios of use of the guard.

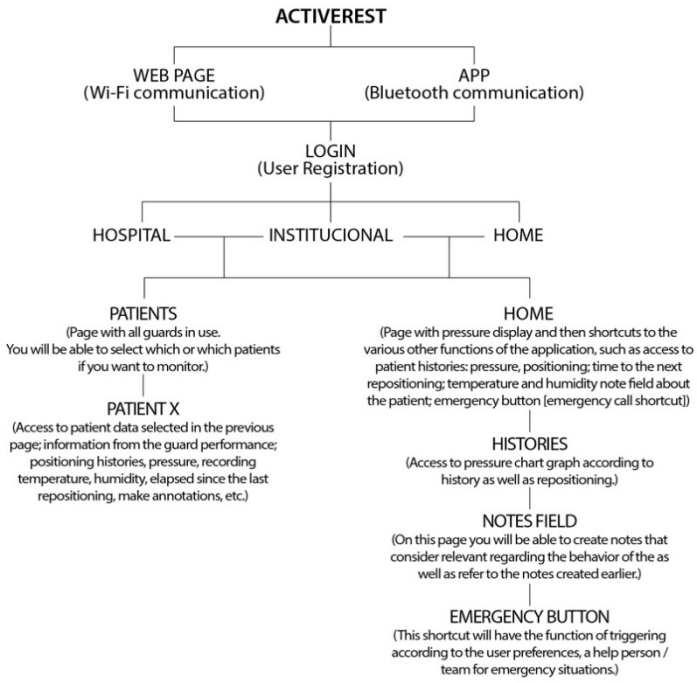


Figure 2: Information Architecture.

6.4. Wireframe and Wireflow

The term wireframe derives from the term waterfall which is defined as a model where the structure process is schematic and sequential, being drawn from top to bottom. This model was described by Winston W. Royce in 1970. With the development of digital editing software, such as Adobe Photoshop, more flexible [27]. According to Brown “Wireframes are" crude "illustrations that show, to a greater or lesser extent, the content of each screen. They are called wireframes because they are usually rendered with simple lines, not very elaborate designs. They illustrate, among other things, what kind of information will be most important on each screen” [28].

From the developed wireframes, it was possible to create the wireflow, that is, the flowchart that will act through the drawn screens. This step made it possible to identify the effectiveness of fluidity in navigating the related functions.

6.5. Visual identity and design guidelines

Regarding this project, the visual identity was developed to act in the mobile application and derived from the logo previously developed for it, and the main components derived from the logo used for the deployment of identity in the app were the colors and minimalist language.

According to the Interaction Design Foundation, design guidelines can be defined as: “sets of recommendations for good design practice. They are intended to provide clear instructions for designers and developers on how to adopt specific principles such as intuition, learning ability, efficiency, and consistency. Rather than dictating conventions, design guidelines provide helpful advice on how to achieve a design principle that can be platform-specific or cross-platform.” (Interaction Design Foundation, 2017). This definition can be understood as a set of good practices to be adopted in the development of design projects. It is not mandatory to use them, but only advisable, to improve the harmony of language for the use [29].

In the user interface field, Google and Apple are the largest guidelines developers, this is because they have the two largest mobile operating systems, respectively, Android and iOS. According to IDC consultant, the two best-selling smartphone brands in Portugal in 2017 were Huawei and Samsung, both using the Android operating system. Nevertheless, a study by Marktest in November 2017 indicated that the sites audited by the Android group had a 30.8% access percentage, representing leadership over other systems [30]. Thus, the most used operating system in Portugal is Android, which is why the app was developed according to the guidelines established by Google for the design of mobile applications.

Google created Material Design, defining it as "a visual language that synthesizes the classic principles of good design with technological and scientific innovation." Thus, this concept can be understood as a set of guidelines that can be used for software development, directed mainly to the Android operating system. On the material.io website, the company offers several packages designed, based on scientific studies, in the concept of guidelines for the user interface. Also defined by UI Kits, this type of content is not required but has a high impact on mobile application development.

In summary, some of the advantages of using these UI Kits are that they are developed based on studies that seek the most satisfactory user possible interface and time optimization. First, on the part of the designer who can use this time in other components of the app design, and secondly on the part of the user, who can create a higher rate of application receptivity, as many components are known by him.

It is also possible to observe in these UI Kits the presence of several recommendations for elderly interface design as well as Nielsen's heuristics. Some examples as simple and straightforward language; visual noise reduction; use of non-serif typography; minimalist design; consistency and predictability in navigation; use of standards among others are inserted in the development of this material. Thus, the use of Google's UI Kit for ActiveRest application development was then defined as it had the interface developed for the Android operating system and these guidelines already fit the recommendations for designing applications for seniors as well as the use of parameters to be used for the adult audience. The images below demonstrate the design of some application screens using Google's UIKit.

6.6. ActiveRest App

After analyzing all the research carried out within all the necessary components for the development of a mobile application that responds to the adult and elderly public in the same digital environment, the application was designed. For its development, the recommendations for interface design for the elderly as well as Nielsen's heuristics were used:

- Color and Icons: The predominant color (# 00AAAD) used is the same as the project logo. For better contrast, black and gray were used, as well as the use of white spaces to prevent visual noise, choosing the minimalist language. The settings can be changed the app theme to a higher contrast mode as well as change the font size. In Figure 3 can be seen the color option used and the option to change the font size. The icons are meant to be easily recognizable, thin lines were avoided, not to disturb the user's perception.

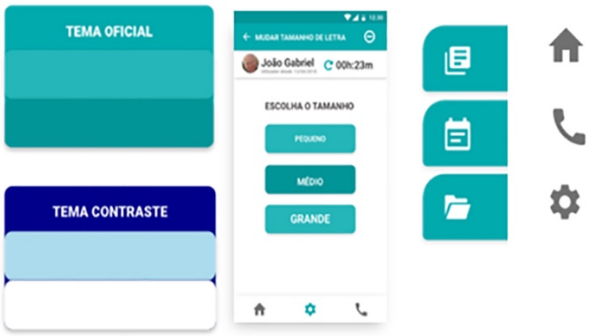


Figure 3: App Colors | Option to change font size | App icons.

- Home and Repositioning Time: on the homepage are the shortcuts to the most important functionalities of the application (see Figure 4), as well as the emphasis of the body pressure map and repositioning time that is present in most of the pages of the app for better control of the user.

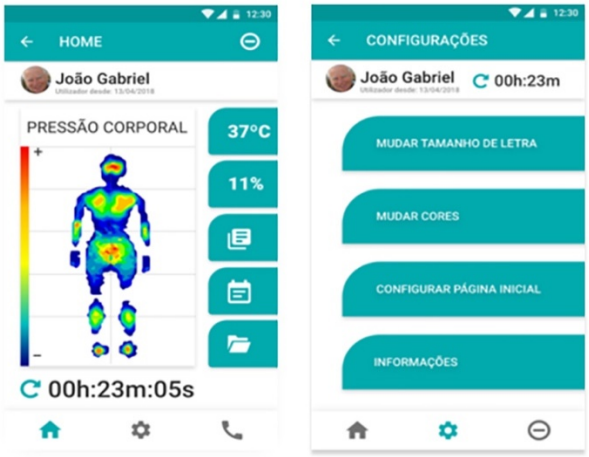


Figure 4: HOME-Domestic | SETTINGS-Hospital / Institutional screen (in the upper right corner is the timer).

- Emergency Button: was developed so that the user performs the minimum of actions to complete the task. We used windows to confirm the fulfilment of emergency requests to avoid unnecessary requests. These windows are positioned closer to the bottom of the interface so that the user can finalize the task with the least possible actions (see Figure 5).

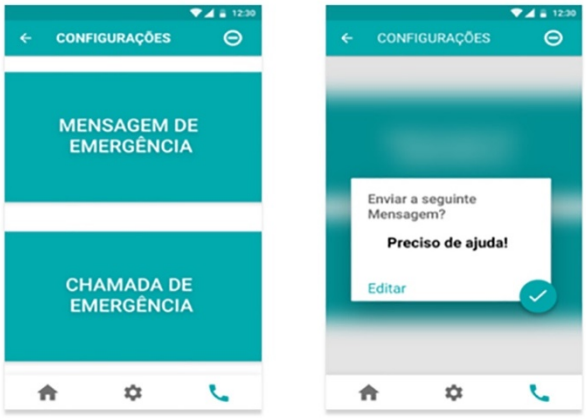


Figure 5: EMERGENCY BUTTON screen and Home confirmation popUp.

- HOME version and HOSPITAL / INSTITUTIONAL version: The two versions of the app are similar in several respects, differing only in the following, the login mode that in the home version is done through the user and password credentials, and in the hospital / institutional version is done through the access PIN, the latter being the most common mode in hospital software; Emergency Button functionality is present only in the home version because it is not useful in a hospital / institutional setting.

6.7. Prototype

The prototypes can be tested in the following links:

- DOMESTIC: https://marvelapp.com/ed2cbie
- HOSPITAL (Institutional): https://marvelapp.com/4id38ji

7. Wireless Communication System

The Bluetooth Low Energy (BLE) communication protocol was selected to enable a simple and standard way to establish between the mattress guard the developed mobile App. For each mattress guard, there will be a local electronic system (comprised of a printed circuit board with a processor, storage unit, and a BLE module), which will manage the power supply of the different electric components and obtain a signal input from them (namely, the sensors and actuators embedded on the mattress guard).

Control of the sensors and actuators is achieved through a microcontroller, which also stores the sensor information in the local storage unit. When the App establishes a connection with the mattress guard, by pairing the mobile device with the BLE module in the local mattress guard controller, data will start being transferred by Bluetooth and become available to the user. Conversely, whenever high pressure loci are detected above the defined threshold, the device will enable one or several of the actuators (according to a protocol developed within the ActiveRest project) in order to eliminate that pressure peak.

Figure 6 shows a schematic of the communication model between the mattress guard and its control and management electronic system, as well as the digital unit to where the data are sent by Bluetooth, becoming accessible to the user.

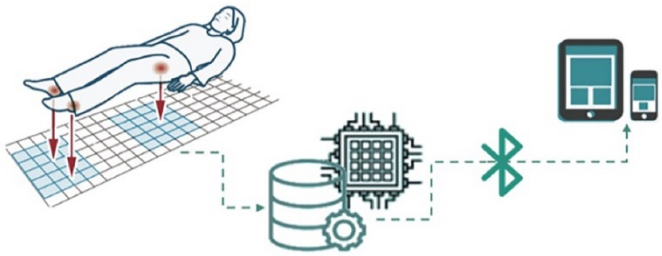


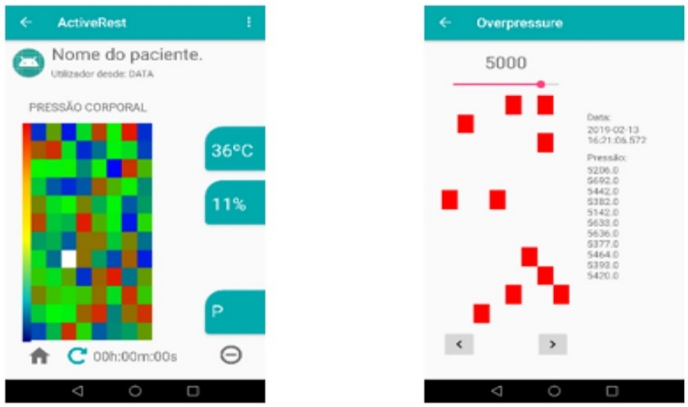
Figure 6: Scheme of the data communication model.

8. Android App Implementation

In order to validate the graphic model of the developed App, a prototype version was created, integrating the core features that can be tested by users and that are vital to the underlying purpose. These include the wireless communication between App and the mattress guard, and are essentially dependent on the hardware that acquires the signal and stores it for further processing. This App was developed using Android Studio version 3.1. It was compiled and further tested using a mobile phone running Android 7.0, and featuring current standard hardware, including the required Bluetooth communications module.

The App initializes with the project logo. After the start-up screen, the user is presented with the first interactive screen, that displays data received from the mattress guard. This main screen intends to provide almost all relevant information the user requires for the purpose of monitoring the bedridden patient, and includes a graphical representation of the pressure mapping. In order to communicate with the mattress guard, the App employs the low energy Bluetooth protocol, using a pre-defined MAC address (configurable by the user), thus receiving data from the external on a regular time interval. The pressure data is processed into a color scale, which highlights specific pressure loci. The employed color scale goes from white (which stands for locations with pressure below a minimum setting, and which are not relevant), passing through a blue gradient (relatively low, but relevant pressure, usually implying the presence of a body), up to the highest pressure values, in red gradient (implying the pressure in that location has exceeded the threshold pressure for potential formation of ulcers). For every sensor present in the mattress guard, the value provided to the App is mapped onto this color scale and rendered in the two-dimensional grid that represents the physical device (see Figure 7.a). It was possible to use a simple and regular two-dimensional grid because all the sensors are spaced equally in the developed mattress guard, but this could be easily adjusted to any other physical configuration of the sensors.

Whenever the local pressure value exceeds the user-defined threshold (for which the project conducted a literature review and collected information from medical staff), it will induce an alarm, which can then trigger both (or either) a communication to a caregiver by text messaging, and a local notification / pop-up (with an audible trigger). Also, in order to allow analysis of the history of previous high pressure loci alerts, whenever the value exceeds the threshold it is stored in a list, which can then be accessed through a specific App page. This second App page, with the history of critical values, is shown in Figure 7.b. Note that in this view only the highlighted pressure loci are usually represented, but the user can quickly adjust the mapping threshold, for example to see how the overall pressure map evolved between alarm intervals. The pressure information is stored in a simple text file which later can be extracted from the device and transmitted to a caregiver or a physician.



(a) (b)

Figure 7: Android App screenshots: (a) the main page where the data from the pressure mapping, and the temperature and humidity values are shown, as well as the repositioning timer, and (b) the page with the history of critical pressures that exceed the established threshold.

Last, the developed App includes a timer that send a notification to the caregiver after the specified interval. This is very useful for the regular repositioning of the patient which is a routine activity that needs to be performed on bedridden patients. The mattress guard does not intent to replace the standard repositioning protocols, but to complement them by identifying whenever a more urgent action is required, and thus, the timer was a feature request by all the professionals that were inquired during the early specification stage.

9. Usability tests

The usability test is one of the most important phases of the project, it is through it that the user will contact the product and will bring the designer answers regarding the usability and effectiveness of the developed interface. Loranger states that the main benefit is identifying design efficiency as well as the actual needs of users [31].

Nielsen states that five participants are the ideal number for taking a test and that, for best results, they should be short-lived and made up of simple tasks [32]. Usability testing can be developed early in the project, except when the product is for senior audiences. In this case, the prototype to be tested must be of high fidelity, because, due to the aging of cognitive abilities, this audience has a lower degree of abstraction and imaginary component [26, 27].

Loranger highlights nine important criteria to consider when preparing a usability test [33, 34]. They are: (1) Define the purpose of the test; (2) Define the format and scenario of the test (lab or field / moderate or unmoderated / face-to-face or remote); (3) Number of users; (4) Ensure that participants are suitable for the intended test; (5) Create tasks corresponding to the purpose of the study; (6) Perform a test run; (7) Collect metrics; (8) Create a script for the test; (9) Conduct the team test consisting of a moderator (who conducts the test) and an observer (who observes and takes notes).

During the test, it is essential to pay attention to the mode of communication with the participant so that the influence on the final action is as small as possible. Nielsen states that the Thinking Aloud method is the most important tool in a usability testusabilidade [35], asking the participant to say what he is thinking while performing the task is an important resource for assessing how the user interprets the interface. Interventions are discouraged when the participant is performing a task, questions should be unbiased to avoid naming interface elements and where possible answer the participant's questions with other questions and these should be exempt leading him to the conclusion by himself [24],[28]-[31].

After the completion of the test it will be through the analysis of the collected data that the success rates of each tested functionality will be found and, if necessary, redesigned to better fulfill the interface objective.

In the present project the usability tests were performed only with the interface developed to work in an institutional / hospital environment, were directed to three groups of participants, the first group was of eight nurses, the second two doctors, and the third five direct action assistants. It was decided to test the prototype in high fidelity in order to facilitate the participants' interaction with the product. A guide was developed for the documentation of the test, which contained a detailed description of each task to be requested for the participant, as well as areas for annotation of the responses and their behavior. An app qualitative assessment questionnaire was also created for each test participant to complete at the end.

For data analysis, the profession and age range of each user was taken into account. The results indicated that the 31-40 age group had a higher success rate in using the interface. In terms of profession, nurses and physicians were the ones that represented the highest success rate. It is noteworthy that according to the questionnaire performed at the end of each test, these two classes were those that showed the greatest interest in mobile application adherence. It was identified that the questionnaire will not be the most appropriate tool to be used in usability tests, due to the fact that it presents different results from those collected in the monitoring guide, that is, the user's opinion did not always coincide with their own rate. of success in accomplishing each task. Results are shown in Figures 8 and 9.

With an 85% success rate, it was concluded that the HOSPITAL/ INSTITUTIONAL interface achieves the proposed objective. However, for better use of the app it is best to have the two features with a lower success rate (change timer and exit the application) be changed as these are key functions in using the app.

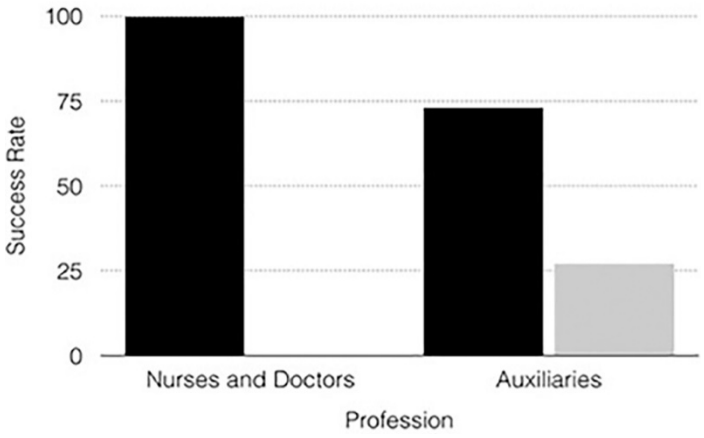


Figure 8: Profession Success Rate Chart; black: success rate; grey: failure rate. The graph comes from usability tests.

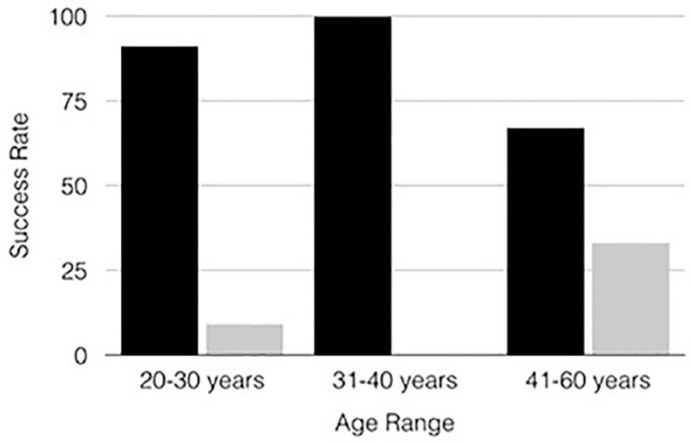


Figure 9: Age Success Rate Chart; black: success rate; grey failure rate. The graph comes from usability tests.

10. Conclusion

Pressure ulcers is a serious problem that significantly affects a patient's quality of life. This problem occurs mostly in the elderly,

with a high loss of physical health. Whether for economic or welfare issues, solutions focused on health care for the elderly are important preventive measures for the well-being of a society in which the average life expectancy of people tends to increase.

ActiveRest emerges as a solution focused on the prevention of UPs, through a textile shelter and a mobile support application. This project aims to offer patients (and their caregivers) a relevant contribution to their quality of life.

The present work had as main objective the development of the digital interface to work together with the 'ActiveRest' textile mattress guard. For the study it was necessary to examine the current panorama of digital technological investment in health with emphasis on the Portuguese market. In this way, it was possible to identify the current reality of health institutions as well as the existing solutions in the market to respond to this problem. The analysis of the behavior of the target public (adult and senior) was fundamental to understand what the real functionalities that these users need. It was also essential to analyze the solutions already on the market, these were selected according to similar functionalities that would be present in the ActiveRest project.

The application resulting from the investigation of this project, adapted with two versions to be used in a hospital environment and another in a home environment, even with the objective of working together with textile protection, is able to independently present support for prevention pressure ulcers. For example, the application's timer functionality provides the user with better control and optimization of time and, if there is a possible failure in blocking the application, it will continue to work. Features like Diary, Help Button and Reminders also do not require the backup feature. For a better use of the application, it will always be necessary to use textile protection, however, these components demonstrate that the mobile application remains operational and functional, even without using it.

11. Acknowledgements

Sandra Ventura (Texteis Penedo), Graça Bonifácio (CITEVE), José Casquilho (ComfortKeepers), Joana Fonseca (CENTI), Joana Almeida (CENTI), Miguel Ribeiro (CENTI), Pedro Pereira (CENTI), and Ricardo Carvalho (CENTI), for their contributions to the ActiveRest project. This research was supported by FEDER funds through the COMPETE 2020 Programme under project ActiveRest (project 18011 of the 33/SI/2015 call).

References

[1] R. Suzman, J. R. Beard, T. Boerma, and S. Chatterji, 'Health in an ageing world - What do we know?', *The Lancet*, vol. 385, no. 9967. Lancet Publishing Group, pp. 484–486, 07-Feb-2015.

[2] E. instituto nacional de PORDATA, 'Esperança de vida aos 65 anos: total e por sexo (base: triénio a partir de 2001)', *PORDATA*, 2014. .

[3] J. Nielsen, 'Middle-Aged Users' Declining Web Performance', 2008.

[4] National Pressure Ulcer Advisory Panel *et al.*, *Prevenção e Tratamento de Úlceras por Pressão: Guia de Consulta Rápida*, vol. 2, no. 1. 2014.

[5] H. Martins, 'Transformação digital e inovação na saúde em debate no "Portugal eHealth Summit"', *PÚBLICO*, 19-Mar-2018.

[6] Serviço Nacional de Saúde, 'Receita Sem Papel - SPMS', 2018. [Online]. Available: http://spms.min-saude.pt/product/receita-sem-papel/. [Accessed: 09-Sep-2018].

[7] SNS, 'Receita Sem Papel – Monitorização'. [Online]. Available: https://www.sns.gov.pt/monitorizacao-do-sns/receita-sem-papel/. [Accessed: 09-Sep-2018].

[8] H. Martins, 'SNS digital e as condições da transformação digital na saúde', 2018.

[9] A Enfermagem e as Leis, 'MySNS Carteira: Aplicação móvel do SNS lidera na categoria de saúde e fitness | A Enfermagem e as Leis', 2018. [Online]. Available: http://www.aenfermagemeasleis.pt/2018/05/09/mysns-carteira-aplicacao-movel-do-sns-lidera-na-categoria-de-saude-e-fitness/. [Accessed: 09-Sep-2018].

[10] G. Lins, 'MySNS Carteira (SPMS) | Guga Lins UX/UI Designer'. [Online]. Available: https://www.gugalins.com/mysns-carteira/. [Accessed: 10-Dec-2019].

[11] Innovatemedtec, 'mhealth - what is mhealth?', 2018. [Online]. Available: https://innovatemedtec.com/digital-health/mhealth. [Accessed: 09-Sep-2018].

[12] C. Duque, J. Mamede, and L. Morgado, 'Iniciativas de mHealth em Portugal mHealth initiatives in Portugal', 2017.

[13] Statista, 'Annual number of mobile app downloads worldwide 2022 | Statistic', 2018.

[14] Statista, 'Global digital health CAGR by major segment 2015-2020 | Statistic', 2018. [Online]. Available: https://www.statista.com/statistics/387875/forecast-cagr-of-worldwide-digital-health-market-by-segment/. [Accessed: 13-Sep-2018].

[15] J. M. Jakicic *et al.*, 'Effect of Wearable Technology Combined With a Lifestyle Intervention on Long-term Weight Loss', *JAMA*, vol. 316, no. 11, p. 1161, Sep. 2016.

[16] K. M. Diaz *et al.*, 'Fitbit®: An accurate and reliable device for wireless physical activity tracking.', *Int. J. Cardiol.*, vol. 185, pp. 138–40, Apr. 2015.

[17] Deloitte, '2016 Technology Fast 500 ™ Ranking Recognizing growth', 2016.

[18] Ministério da Saúde, 'Relatório anual acesso a cuidados de saúde nos estabelecimentos do sns e entidades convencionadas', 2018.

[19] OECD, *Health at a Glance 2017*. OECD, 2017.

[20] U.S. Department of Veterans Affairs, 'Resource App', no. March, 2017.

[21] XSENSOR, 'XSENSOR Technology - Accurate Pressure Imaging Sensors'. [Online]. Available: https://xsensor.com/. [Accessed: 10-Dec-2019].

[22] uControl, 'uControl'. [Online]. Available: https://www.unitron.com/content/unitron/pt/pt_PT/professional/hearing-solutions/accessories/hearing-instrument-accessories/ucontrol.html. [Accessed: 10-Dec-2019].

[23] Pressure Ulcer Prevention (PUP) |, 'Pressure Ulcer Prevention (PUP) |'. [Online]. Available: https://medicinegov.org/pressure-ulcer-prevention-pup/. [Accessed: 10-Dec-2019].

[24] BodiTrak, 'BodiTrak – Wireless Bed Pressure Mapping System'. [Online]. Available: https://www.boditrak.com/products/medical/bed-system.php. [Accessed: 10-Dec-2019].

[25] Internetmedicine, 'The Lifelens Project: Smartphone App Diagnoses Malaria from Drop of Blood | InternetMedicine.com', 2014. [Online]. Available: http://internetmedicine.com/2014/04/26/malaria1/. [Accessed: 13-Sep-2018].

[26] J. Nielsen, '10 Usability Heuristics for User Interface Design', *Nielsen Norman Group*, 1995. [Online]. Available: https://www.nngroup.com/articles/ten-usability-heuristics. [Accessed: 12-Dec-2019].

[27] R. S. Wurman, P. Bradford, and G. P. Corp, *Information Architects*. Graphis, 1997.

[28] Timetoast, 'History of wireframing Timeline', *History of wireframing*

Timeline, 2018. [Online]. Available: https://www.timetoast.com/timelines/23833. [Accessed: 28-Aug-2018].

[29] G. L. Strunck, *Como criar identidades visuais para marcas de sucesso.*, 4ª. Rio de Janeiro, 2007.

[30] Interaction Design Foundation, 'What are Design Guidelines? | Interaction Design Foundation', 2017. [Online]. Available: https://www.interaction-design.org/literature/topics/design-guidelines. [Accessed: 06-Sep-2018].

[31] H. Loranger, 'Usability Test, Even When You Know the Answer', 2018. [Online]. Available: https://www.nngroup.com/articles/test-when-you-know-answer/?lm=why-you-only-need-to-test-with-5-users&pt=article. [Accessed: 01-Dec-2019].

[32] J. Nielsen, 'Why You Only Need to Test with 5 Users', *Nielsen Norman Gr.*, 2000.

[33] A. Dickinson, J. Arnott, S. Prior, and A. Dickinson, 'Methods for human-computer interaction research with older people', *Behav. Inf. Technol. is*, vol. 26, no. 4, pp. 343–352.

[34] H. Loranger, 'Checklist for Planning Usability Studies', 2016. [Online]. Available: https://www.nngroup.com/articles/usability-test-checklist/?lm=users-real-data&pt=article. [Accessed: 01-Dec-2019].

[35] J. Nielsen, 'Thinking Aloud: The #1 Usability Tool', 2012. [Online]. Available: https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/. [Accessed: 01-Dec-2019].

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 683-687 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Classification of Timber Load on Trucks

Jan Sikora*, David Fojtík

Department of Control System and Instrumentation, VSB – Technical University of Ostrava, 708 00, Czech Republic

ARTICLE INFO

Article history:
Received: 15 January, 2020
Accepted: 24 March, 2020
Online: 20 April, 2020

Keywords:
Truck
Wood
Neural network
Classification
Image

ABSTRACT

All trucks heading into the paper mill MONDI, Slovakia, have to pass an automatic security check. It controls if storage of its wood load meets all standards of safety. Each truck is scanned by a group of 2D scanners. After that the inspection of timber load is done by a software with use of the data gained by these scanners. The security software is universal for all kinds of storage of timber loads. This article is dedicated to deal with a problem of classification a kind of wood storage on a semi-trailer. The classification is solved by training a convolutional neural network on datasets with recorded trucks of both kinds to learn patterns distinguishing them. The image classification is done with use of images recorded by a set of cameras. By determining a type of storage, it is possible to execute the safety check for a specific type of wood load with better result than the universal check.

1. Introduction

Deep learning as a promising field has been quickly involving during last years. Artificial neural networks, which are cornerstone of this field, have proven to achieve overwhelming results in many disciplines. Computer vision is one of such area, where deep learning techniques are being used for solving variety of tasks. Me and my colleagues were working on a particular computer vision task in the past. Back then, we solved the problem by analytic method with use of standard computer vision functions. Recently we decided to solve the same task by using deep learning techniques and compare the results. These objectives will be described in this paper, which is an extension of work originally presented in proceedings of the International Carpathian Control Conference 2019 [1].

Computer vision belongs to technologies in which deep learning techniques are widely used nowadays. Some of applications are for example image recognition, image processing [2], object detection [3], solving style transfer problem [4]. We cannot omit image classification problem [5], which is the issue of this paper.

Our classification problem is binary, i.e., the developed software must assign one of only two classes to the provided input data. In this particular case we are classifying kind of timber load stored on trucks. The classification should lead to a improve safety check of trucks heading into paper mill MONDI SPC, a.s. in Ruzomberok.

Similar problems of image classification have been solved in the past by applying supervised machine learning algorithms, which is the case of the article focused on classifying agricultural landscapes algorithms [6]. Next article solves the problem of cloud detection, that is binary classification problem, with use of large-scale gaussian processes classifier [7]. Next study shows a classification of heat emitting object with use of Convolutional Neural Network [8]. Network of such kind is also used in this work.

2. Recording Trucks

Before entering the paper mill MONDI SPC, each truck loaded with wood logs has to pass a safety check. For this purpose, a unique gate equipped with a set of 2D lidar scanners and a group of cameras has been developed and installed. Functionality and design of this gate are described in detail in the older article [9].



Figure 1: The scanning gate (left) with detail on a camera (right).

*Corresponding Author: Jan Sikora, Email: jan.sikora@vsb.cz

Software executing the safety inspection of trucks use point cloud that represent a surface of each truck. These point clouds are obtained by the set of scanners while trucks are passing through the gate. The algorithm for safety check is universal for all kinds of timber load. We have an idea that can lead to improve success rate of the inspection, if we use a specific algorithm for relevant kind of timber storage.

But for that, we need to classify kind of wood storage of each truck, before running one of specifics inspection algorithm. This classification is the main goal of both original and this article. For this purpose, we can use the installed cameras, which serves now only for recording and storing an evidence of incorrectly loaded trucks.



Figure 2: Pictures of a truck from all three cameras.

3. Classification with use of Computer Vision Technology

The original article describes how we solved this classification problem with use of analytic functions belonging to the computer vision technology. These functions are described in detail in the book [10]. A final classification result of individual truck was determined by mean results of two classification methods.

3.1. Top View Classification

First method works with images taken by a camera mounted on the top of the scanning gate. At the beginning of the recognition algorithm, the important area of taken image (highlighted by the blue quadrangular on the left part of Figure 3) is converted by perspective transformation into simulation of top view. Then the image is converted into grayscale one, on which the Canny algorithm for finding edges is applied.

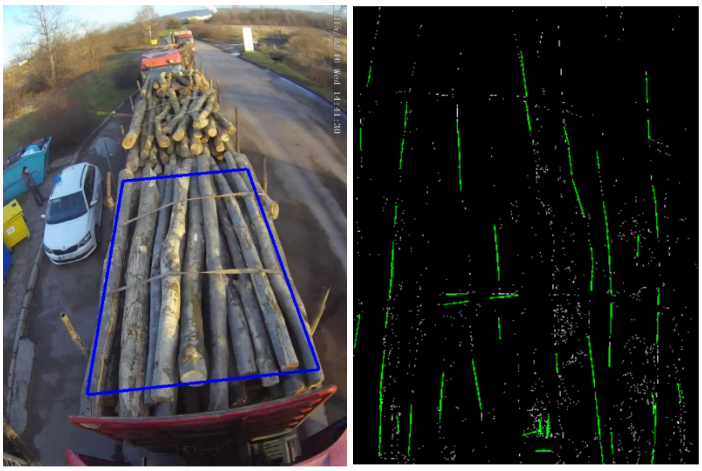


Figure 3: Important area of image (left) and final image with founded lines [1].

Then lines are detected by use of Hough algorithm (see right part of Figure 3). Direction of those lines shows a type of wood storage.

3.2. Side View Classification

This method uses images taken by cameras mounted on both sides of the scanning gate. Principle of this method is to find cross-sections of wood logs in truck's images. When they appear in significant percentage of image's area, then a truck is classified as transversely loaded one.

At first, via perspective transformation, an important part of an image is converted into a side view simulation. The perspective transformation is done due to easier wood logs' cross-sections detection when they transform from shape of ellipses to more or less circular shape.



Figure 4: Important area of image (left) and its perspective transformation into simulation of side view [1].

Then threshold is applied to find cross-sections of wood logs. Pure wood (cross-sections of wood logs) occur in variations of yellow color, from light to dark yellow. Whereas, tree bark (cylindrical surface of wood logs) occur in shades of brown or gray color. This feature is used in image processing to detect cross-sections of wood logs. Pixels belonging to cross-sections of wood logs are detected by applying threshold that define if pixels are within the range of yellow colors in HSV color spectrum (see Figure 5).

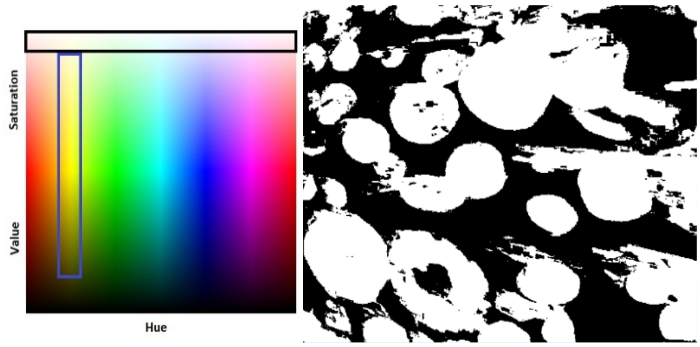


Figure 5: HSV color spectrum with marked ranges of yellow color (left) and threshold of both color ranges merged together on side view image (right)

At first, the algorithm looks for cross-sections of wood logs separately on both images of thresholds of yellow color, and later on merged one.

The erode function is called on the threshold images repeatedly with different number of iterations. At first, the erode function with relatively large amount of iterations is called to split and detect bigger blobs. Afterwards, the detected blobs are subtracted from the threshold original image, and the erode function with less amount of iterations is executed to find small blobs as well.

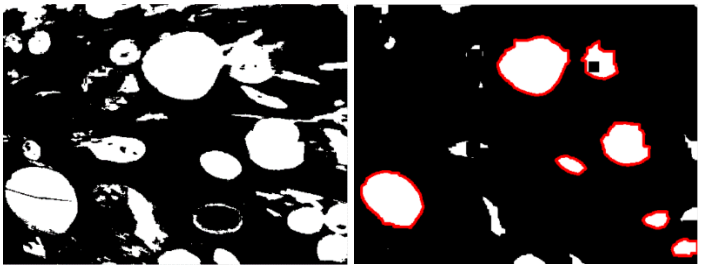


Figure 6: Threshold of one of yellow colour (left) and its blob detection (right)

Detected blobs found on all three threshold images are approximated to circles. Those circles are drawn, into the image of its original perspective transformation, for better visualization (see Figure 7).



Figure 7: Founded cross-sections of wood logs on transversely loaded truck (left) and on longitudinally loaded one (right) [1].

As you can see, cross sections are also detected on images of trucks with longitudinal kind of timber load. But, to classify a truck as a transversely loaded one, areas of the cross-sections must occupy significant part of the image's total area.

4. Classification with Use of Deep Learning Technology

Image classification is one of the problems where deep learning really excels. The high success of neural networks in the field of computer vision is achieved especially thanks to convolutional layers, which works similar as human cortex. [11].

Despite long training time, neural networks are very fast and effective when they are already deployed to solve a specific task. Our image classification problem belongs to family of supervised learning. Neural networks of this group can be very accurate if the training dataset is huge and heterogenous, because in that case they learn a more general pattern [12]. Unfortunately, our dataset is rather small due to the fact, that transversely loaded trucks are extremely rare.

We had pictures of only 412 transversely loaded trucks at disposal. On the other hand, a list with records of longitudinally loaded trucks was massive. Even though, we used only 724 of these to keep moreover balanced datasets of both classes. The input data were divided into three groups. First one was used for training a neural network, second one was dedicated for validating the neural network during training, and last group was reserved for final testing of trained network.

Table 1: Distribution of input data into datasets.

	train	validation	test
longitudinal	406	156	162
transverse	231	89	92

Speed of trucks moving through the scanning gate is between 5 and 10 km/h. The cameras have a frequency of only one record per a second. Each camera takes on average 8 pictures of a truck during its scanning. Since we are using 3 cameras, each truck is represented by 24 pictures on average. The images of trucks are taken with a relatively high resolution of 2048 x 1536 pixels by the left and the right camera. Resolution of the top camera is the same, but swapped in axes, so its value is 1536 x 2048 pixels. All input images are resized to 512 x 512 pixels before entering the neural network. This action makes both evaluation and especially training phase significantly faster. The input images that are fed into the neural network, have three channels, that represent RGB color format.

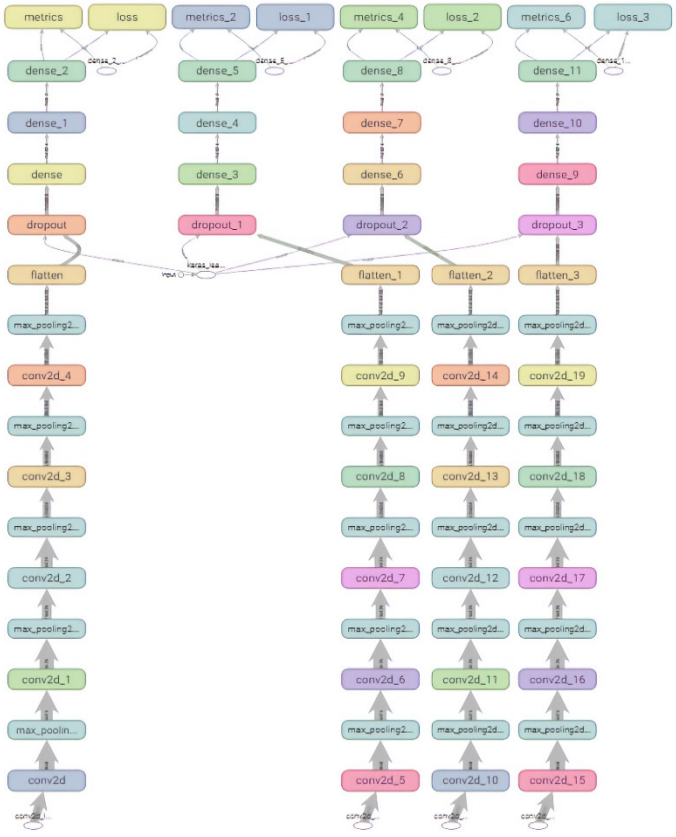


Figure 8: The graph of used architecture.

The architecture of used neural network is based on VGG-19 model [13]. After using this model architecture to our classification dataset, we have done some experiments with the model configuration and slightly change its architecture. Final model is sequential, rather deep and consist of five convolutional layers. Every convolutional layer use ReLU (Rectified Linear Unit) as an activation function [14], and their kernels have a shape of 3 x 3. The number of filters increases with the depth of the model. First convolutional layer has only 32 filters, second one is made up from 64 filters, third and fourth layer have the same amount of 128 filters, and the last convolutional layer consists of 256 filters. After each of these convolutional layers, a max pooling operation, with pooling size of 2 x 2, is executed.

Then the model is flattened, and dropout function with rate of 0.5 is activated. In the end of the model are three densely connected layers. The number of units per layer is decreasing with depth from 1024 via 512 to 1. First two of these densely connected layers use ReLU activation function, whereas the last layer rely on Sigmoid activation function.

The neural network was trained for 120 epochs. The training has been done with RMSprop optimizer with 0.0001 learning rate. Choosing the hyper-parameters of the network have been done based on papers [15, 16]. Accuracy of the network was verifying on the validation dataset during training. As we can see from Figure 8, the validation accuracy stopped to follow the growing tendency of the training accuracy around 100th epoch. After that point the network started to memorize input data rather than finding a better classification pattern.

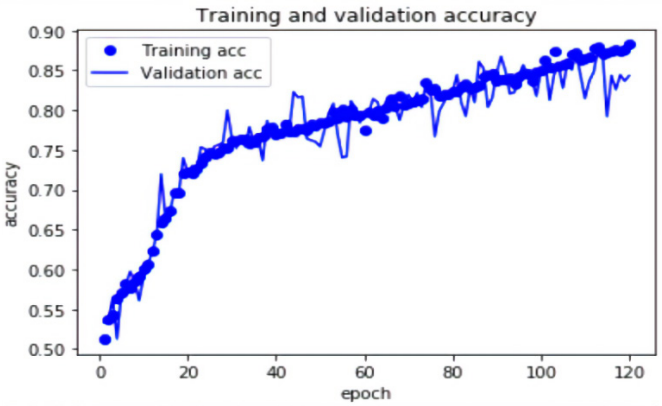


Figure 9: Training the network.

After initial training, a new neural network was trained with only a hundred epochs to avoid overfitting [17]. To combat overfitting, we used the techniques described in [18, 19], which help us to choose the best elements of model dealing with that problem, and set their parameters. These techniques were proven, via comprehensive experiments, to increase the model's classification success rate on new data. This final network was later examined on the testing dataset. The final classification accuracy achieved a result of 84.926 %, which is a great outcome received from our limited training dataset. With use of larger training dataset, we could achieve even greater success rate.

Although we tried to solve this task by training many different neural networks architectures with convolutional layers, the one described in this chapter has the best result.

The classification accuracy of the original paper with analytic approach has only slightly better result, that reached to 85,086 %. While maintaining almost identical classification accuracy, the new method is few times faster and need less computing energy.

5. Conclusion

The classification accuracy achieved by the trained convolutional neural network is 84.92%. This number is accuracy of an individual truck's image. When we classify a truck according to a mean result of all its images, we can reach up to 97.88% success rate.

This classification problem has already been solved in the original paper by analytic methods with use of computer vision technology. Back then we achieved similar result of 98% success rate.

Classification approach of this paper uses artificial neural network, and as such, it is more reliable in bad weather and light conditions. For example, snowing, raining, cloudy weather, or darkness at night, can lead to huge unreliability of analytic approach of the original paper, since it depends on finding yellow colour of wood log cross-sections. On the other hand, actual solution can be reliable even in those bad conditions, when the images of the training dataset include images with these situations.

Even though actual solution has not brought much better result, it is far faster and more energy efficient solution than the original one. The big future potential is by combining these two methods. The determining if a truck is longitudinally or transversely loaded will reach almost a hundred percent certainty by that. Then, we can upgrade the actual software for truck's safety inspection to work for a specific kind of timber load, and by that the safety check will be more precise.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by the European Regional Development Fund in the Research Centre of Advanced Mechatronic Systems project, CZ.02.1.01/0.0/0.0/16_019/0000867 within the Operational Programme Research, Development and Education and the project SP2020/57 Research and Development of Advanced Methods in the Area of Machines and Process Control supported by the Ministry of Education, Youth and Sports.

References

[1] J. Sikora, D. Fojtík, J. Czebe and M. Mihola, "Storage kind recognition of truck's timber load", 2019 20th International Carpathian Control Conference (ICCC), Krakow-Wieliczka, Poland, 2019, pp. 1-4. https://doi.org/10.1109/CarpathianCC.2019.8766049

[2] M. Jurek and R. Wagnerová, "Frequency Filtering of Source Images for LASER Engravers", 2019 20th International Carpathian Control Conference (ICCC), Krakow-Wieliczka, Poland, 2019, pp. 1-5. https://doi.org/10.1109/CarpathianCC.2019.8766050
[3] S. Kanimozhi, G. Gayathri and T. Mala, "Multiple Real-time object identification using Single shot Multi-Box detection", 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019, pp. 1-5. https://doi.org/10.1109/ICCIDS.2019.8862041
[4] L. Gatys, A. Ecker and M. Bethge, "Image Style Transfer Using Convolutional Neural Network", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 2016, pp. 2414-2423. https://doi.org/10.1109/CVPR.2016.265
[5] H. R. Roth et al., "Anatomy-specific classification of medical images using deep convolutional nets", 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), New York, NY, 2015, pp. 101-104. https://doi.org/10.1109/ISBI.2015.7163826
[6] D. C. Duro, S. E. Franklin, M. G. Dubé, "A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery", 2012, Remote Sensing of Environment, vol. 118, pp. 259-272. https://doi.org/10.1016/j.rse.2011.11.020
[7] P. M. Álvarez, A. P. Suay, R. Molina and G. C. Valls, "Remote Sensing Image Classification With Large-Scale Gaussian Processes", 2018, IEEE Transactions on Geoscience and Remote Sensing, vol. 56, pp. 1103-1114. https://doi.org/10.1109/TGRS.2017.2758922
[8] A. D. Algarni, "Efficient Object Detection and Classification of Heat Emitting Objects from Infrared Images Based on Deep Learning", 2020, Multimedia Tools and Applications, pp. 1-24. https://doi.org/10.1007/s11042-020-08616-z
[9] F. David, P. Petr, M. Miroslav and M. Milan, "Scanning of trucks to produce 3D models for analysis of timber loads", 2016 17th International Carpathian Control Conference (ICCC), Tatranska Lomnica, Slovakia, 2016, pp. 194-199. https://doi.org/10.1109/CarpathianCC.2016.7501092
[10] G. R. Bradski, Learning OpenCV, Sebastopol: O'Reilly, 2008.
[11] W. Rawat, Z. Wang, "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review", Neural Computation, **29**(9), 2352-2449, 2017. https://doi.org/10.1162/neco_a_00990
[12] F. Chollet, Deep learning with Python, Manning Publications Co., 2018.
[13] A. Koesdwiady, S. M. Bedawi, C. Ou and F. Karry, "End-to-End Deep Learning for Driver Distraction Recognition", 2017, Image Analysis and Recognition, ICIAR 2017, Lecture Notes in Computer Science, vol. 10317, Springer, Cham. https://doi.org/10.1007/978-3-319-59876-5_2
[14] A. F. Agarap, "Deep Learning using Rectified Linear Units (ReLU)", arxiv 2018, ArXiv: 1803.08375
[15] L. N. Smith, "A disciplined approach to neural network hyper-parameters: Part 1 -- learning rate, batch size, momentum, and weight decay", 2018, ArXiv: 1803.09820
[16] H. Wu, Q. Liu and X. Liu, "A Review on Deep Learning Approaches to Image Classification and Object Segmentation", 2019, CMC-Computers, Materials & Continua, 60(2), 575–597. https://doi.org/10.32604/cmc.2019.03595
[17] H. Zhang, L. Zhang and Y. Jiang, "Overfitting and Underfitting Analysis for Deep Learning Based End-to-end Communication Systems", 2019 11th International Conference on Wireless Communications and Signal Processing (WCSP), Xi'an, China, 2019, pp. 1-6. https://doi.org/10.1109/WCSP.2019.8927876
[18] A. F. Kamara, E. Chen, Q. Liu and Z. Pan, "Combining contextual neural networks for time series classification", 2020, Neurocomputing, vol. 384. pp. 57-66. https://doi.org/10.1016/j.neucom.2019.10.113
[19] L. Pan, C. Li, S. Pouyanfar, R. Chen and Y. Zhou, "A Novel Combinational Convolutional Neural Network for Automatic Food-Ingredient Classification", 2020, Computers, Materials & Continua, vol. 62, pp. 731-746. https://doi.org/10.32604/cmc.2020.06508

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 688-696 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

ROS Based Multimode Control of Wheeled Robot

Rajesh Kannan Megalingam*, Santosh Tantravahi, Hemanth Sai Surya Kumar Tammana, Nagasai Thokala, Hari Sudarshan Rahul Puram, Naveen Samudrala

Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India

ARTICLEINFO

Article history:
Received: 15 January, 2020
Accepted: 25 March, 2020
Online: 20 April, 2020

Keywords:
Mobile robot
Python
Virtual network computing
Joystick
IoT
Android app
ROS
Multi-mode control
Speech
IP Address
Socket
Server - Client Protocols

ABSTRACT

This research work mainly presents the design and development of a small-scaled wheeled robot, which can be controlled using multiple controlling interfaces using some new technological trends. Raspberry Pi 3 as the main controller, Python as the programming language integrated with the Robot Operating System (ROS) and Virtual Network Computing (VNC) for screen sharing are utilized in order to control the robot. Laptop Keys, Graphical User Interface (GUI) on PC, Virtual Network Computing on smart phones, Joystick connected to Laptop using python integrated with Robot Operating System, Voice commanding through an android app integrated with python and ROS and Internet of Things using a customized webpage with Node MCU as the controller are the dominant control interfaces. A Local Area Network is established in order to communicate with the micro-controllers wirelessly. The robot can be controlled with multiple interfaces by using its corresponding commands. The robot is made to survive in five different paths, controlled by all the six methods and the time for each response obtained was appraised and illustrated properly.

1. Introduction

Mobile robots are becoming increasingly popular across various sectors. They are used to assist with work processes and even accomplish tasks that are impossible, unreachable or dangerous for humans. Study of these robots had become an emerging field in the research world. Autonomous mobile robots (which can navigate themselves through an environment), guided robots (which travel through a pre-defined navigation route) and finally controllable robots (which travel according to the command received from a device). There are several types of mobile robots, which are controlled by different methods, depending on the need and operational conditions in the application environment. These robots are made resistant to harsh environments and are also available in several sizes too. Mobile robots had also introduced new trends in research.

The mobile robot, along with its control interfaces, is a simple device, made up of readily available materials. Raspberry Pi 3, which supports wireless communications, was selected as the single-board computer. PYTHON IDLE had been used as programming software to build keyboard and GUI control methods. The programs used for implementing the control methods were written in Python and were ROS integrated to implement joystick and voice control methods. This project also uses a Virtual Network Computing (VNC) viewer for smart phone purposes.

Six different methods were used to move the robot in desired direction: laptop PC, smart phone, GUI, joystick, voice command, IoT. The experimental results of three of these methods were already presented in [1] and in this research work, we added three more methods as extension of it. In [2] the voice command of eight languages is used for security system where as the app designed in this research work can take input command of many languages and can control the bot accordingly similar to that of [3]. Mobile control had been achieved using the VNC viewer by reducing its screen encoding time using the technique in [4]. IoT

* Dr. Rajesh Kannan Megalingam
Director, Humanitarian Technology (HuT) Labs, Amrita Vishwa Vidyapeetham, Kollam, Kerala, India, PIN: 690525.
Email address: rajeshkannan@ieee.org, Contact No: +91 9496120900

control method is similar to [5] but instead of an WIFI module and raspberry pi, we used NodeMcu which can both receive WIFI signal and control the robot. GUI and the laptop keys control were implemented by using raspberry pi as both microprocessor and microcontroller, where additionally Arduino in [6], and DSP in [7] were used for controlling the robot. This research work identifies the best method out of six selected modes of control. The best is determined based on the time difference between the time input is given and the robot finishes execution. This is to determine the speed efficiency of the robot, which is a major constraint in most of the situations. These tests conducted can be used to determine the speed with which these commands are transmitted and the processing speed of the controller board.

2. Motive

In recent times mobile robots have become more commonplace in various commercial and industrial applications. These robots are useful in reaching the places that are beyond human reach. Their uses are wide ranging. Primary motive behind the implementation of these mobile robots is to improve speed and efficiency. These robots when equipped with several types of sensors can improve functionality over wide range of applications. The robot can either move in an environment guided by the sensory inputs or be made to follow a specified trajectory. Usually, a terminal (computer user interface used to give input) is essential for controlling the bot. In situations where time is major constraint such as rescue robots, we need to send more commands to control the robot, our aim is to identify the best method for controlling such robots. Experimental results for three methods of control i.e. using Laptop, GUI, Mobile were tabulated before and are now extended by adding three more methods i.e. Joystick, Voice, IOT.

3. Related works

Condry and Nelson [8] presented a model that integrates control system gateways with smart IoT devices, using real-time challenge-response authentication for assured control operations. Agrawal et al. [9] offered a training method and algorithm, using a force-feedback joystick, with an “assist-as-needed” model for training vehicle drivers. [10] proposed a software structure, ASSOCIATES-ROS, for development of unresponsive, collaborative, multi-robot systems with copious software resources Shahzad [11] enlarged the Pymote structure to simulate packet-level performance. The extensions included radio propagation, energy consumption, and mobility. [12] discussed a new form of augmentative and alternative communication (AAC) device, for people with severe speech impairment. [13] proposed a method, to integrate the standard and throat microphone signals, for robust voice identification in clamorous environments. Kim’s team [14] proffered a novel approach to an auditory prototype, for robust speech identification in clamorous environments. [15] presented a book and straight forward hand gesture identification method, used in the mending of people who have flexibility issues, specifically stroke patients and victims with spinal cord injury (SCI).

The authors of [16] explained the simulation of mobile robot using three stage torque control. Author delineated [17] implementation and assessed viability and execution of an embedded application of the Convolutional Neural Network (CNN) algorithm on the Raspberry Pi 3 computer. The CNN-embedded algorithm classifies and quantifies dissimilarities among the frames. The convolutional neural network (CNN) algorithm which they used to detect faults in SHM works by taking an input image, assign importance to various aspects/objects in the image and be able to differentiate one from other and counts these variations as faults, which may not be true always. This algorithm may give a count more than the actual number of faults. [18] describes the execution and working principles of transportable, flimsy user-interfaces to approach implanted distant home-systems and home-networks. Authors in [19] proposed RemoteUI, which performs more efficiently compared to VNC in terms of bandwidth.

Mobile robot with single control interface is presented in [20]. Stabilization and tracking of a robot using a predictive controller with a time varying system is proposed in [21]. Isolated control interface, push-button control based movable robots in an untold domain using wireless sensor technology is presented in the research work [22]. [23] concentrates on avoiding singularity issues rather than the control interface, in an over-stimulated, artificial poly directional, wheeled movable robot [24] used the robot operating system platform for planning motion of poly directional wheeled movable robot instead of the control interface. Authors of [25] explained a novel method to control an orthotic arm using the EOG signals and also proposed a method by which a second person could control it. A low cost auto-navigated wheelchair is designed by authors of [26] using fixed path navigation with a Raspberry Pi, Arduino and rotation encoders. It works on the principle of Distance estimation.

4. Architecture of the system

The system architecture, as shown in Figure 1, is comprised of three major components: the command block, controller block, and the mobility block. The command block is constituted of six interface units: keyboard, smart phone, GUI, joystick, voice command, IoT. Keyboard is the first control interface. The VNC viewer installed on the smartphone is used as the second control method. The third robot mobility control method is the GUI, installed on any PC; next is a joystick connected to a computer or laptop. Voice commands, sent through a mobile app and the sixth method is through IoT (Internet of Things) using a Node MCU (Single board Microcontroller) with the commands coming from a smart phone. The Wi-Fi communication interface is used for transmitting commands and controlling the robot. Single-board computer i.e. Raspberry Pi 3, a Node MCU interfaced with Wi-Fi and motor drivers make the controller block. The mobility block is comprised of motor drivers, the wheels and the motors. Python is the language used to send commands from microcontroller to

motor driver and in some of the methods, we used python programs integrated with ROS, to send commands.

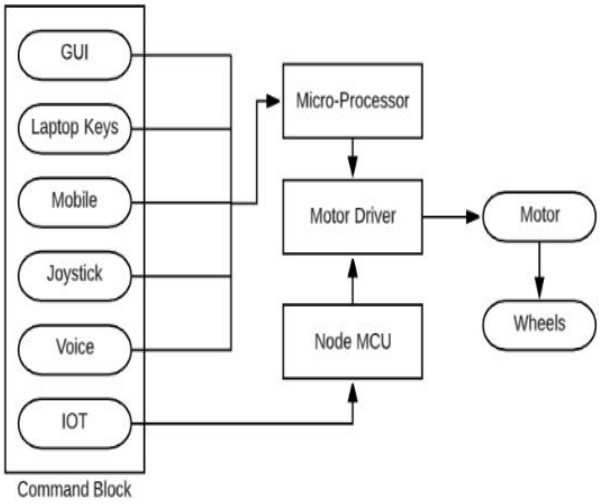


Figure 1: Robot Architecture

4.1 Command Block

Graphical User Interface (GUI) allows the control of robot through simple graphics. The GUI was developed using Python language, five different graphical buttons were created which helps in controlling the robot: up, down, left, right and stop. The robot's front and back motion is controlled by the up and down buttons. The left turn, right turn and stop abilities of the robot is controlled with the help of left, right and stop button respectively.

VNC Viewer app is one of the means to control a wheeled bot, by a smart mobile phone through Wi-Fi communications. VNC viewer is used to control another computer by displaying its visual desktop on host computer connected over a network. Ensure that you have been given permission to remotely connect to the host computer. Once authentication is achieved, the VNC Viewer shows desktop of host computer in a new window. This is similar to the sharing desktop sharing of PC and controlling robot. VNC adapts itself automatically and dynamically to varying conditions, including differing screen contents and network bandwidths. VNC is also platform independent i.e. a Windows system can control a Linux system.

The keyboard of a laptop is used to send commands to the robot and control it to move in the desired direction. We used I, J, K, L and S keys in laptop to control the robot in various directions. Keys i, k, j, l and s are used to move the robot in forward, back, left, right directions and to stop the bot respectively.

A joystick is a input device which provides better control over the robot. It is easier to use by a human operator and steer a bot. Joystick often has one or more pushbuttons, called switches, whose positions can be monitored by the computer. The working of a joystick is based on the principle of conversion of physical movement into a digital signal, which when accepted by the device being controlled, produces the same results on the screen when connected to a computer.

4.2 Mobility Block

The mobility block is constituted of motor drivers, wheels and DC motors. Because of speed efficiency DC motors were better than stepper motors and also DC motors were low weight, small size, and minimum cost. Speed of a DC motor can be controlled by controlled by varying the supply voltage of the controller. In the context of this paper's main objective, the robot should be as compact and lightweight as possible.

4.3 Controller Block

The controller block's main components are a Raspberry Pi 3 computer and Node MCU. One of the six interfaces were used to send the commands. The controller block receives signal through WIFI, decodes these commands and generates control signal to the required modules to control the bot. The direction in which the robot moves is also being decided by the controller block depending on the received command.

5. Modelling the robot

The modeling of the robotic movements was carried out through three cases. The cases include - normal plane ground, robot going up an inclined plane and last, the robot going down the inclined plane. This kind of functional modeling offers an insight into the working of the robot in different terrains.

For all the equations presented here the following symbols are used and the meaning are as follows

N-Normal Force, Unit: Newton

m-Mass of robot, Unit: Kg

g-Acceleration due to gravity, Unit:m/s^2

fs-Static force of friction,Unit:Newton

μ_s -Coefficient of static friction

i_d -Current drawn, Unit: Ampere

τ -Torque, Unit: Newton meter

K_T -Torque constant for a DC motor, Unit: Newton meter/Amp

r-Radius of wheel

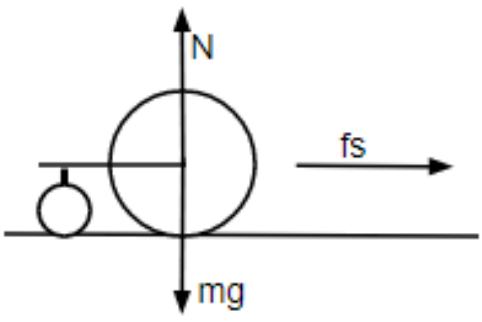


Figure 2: Robot on a flat surface

Considering the case of inclination angle zero i.e. plane ground, as shown in Figure 2,

$$N=\text{mg} \quad (1)$$

The force required to overcome friction is

$$F=\mu_s*N \quad (2)$$

$$F=m*g*\mu_s \quad (3)$$

The torque for the wheels comes from the DC motors

$$\tau=r*F \quad (4)$$

$$\tau=r*m*g*\mu_s \quad (5)$$

For a dc motor with constant input voltage

$$\tau=K_T*i_d \quad (6)$$

$$K_T*i_d=m*g*\mu_s \quad (7)$$

$$i_d=m*g*\mu_s/K_T \quad (8)$$

Here, i_d is the minimum current requirement for the bot.

Similarly, considering the cases of inclination, where the angle of inclination is θ, as shown in Figure 3.

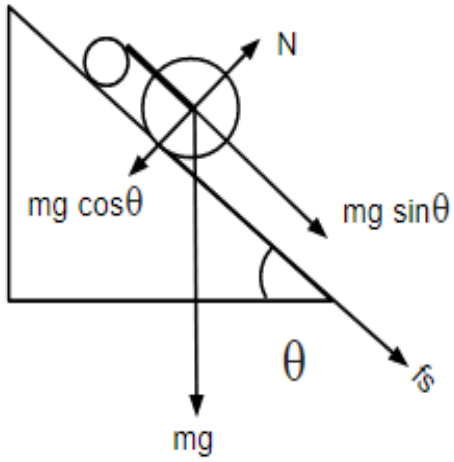


Figure 3: Robot moving up an inclined plane

$$N=m*g*cos\theta \quad (9)$$

Assuming the frictional force, between the inclined plane and the tires of the robot, is f_s.

The force required to overcome friction is

$$F=(\mu_s*N)+(m*g*sin\theta) \quad (10)$$

$$F=(m*g*\mu_s*cos\theta)+(m*g*sin\theta) \quad (11)$$

$$F=r*m*g(cos\theta*\mu_s+sin\theta) \quad (12)$$

The torque for the wheels comes from the motors connected.

$$\tau=r*F \quad (13)$$

$$\tau=r*m*g(cos\theta*\mu_s+sin\theta) \quad (14)$$

For a dc motor with constant input voltage

$$\tau=K_T*i_d \quad (15)$$

$$K_T*i_d=r*m*g(cos\theta*\mu_s+sin\theta) \quad (16)$$

$$i_d=r*m*g(cos\theta*\mu_s+sin\theta)/t \quad (17)$$

Here, i_d is the minimum current'

Similar to the above case, considering the inclined plane with an inclination angle θ and the robot moving down the plane as shown in Figure 4.

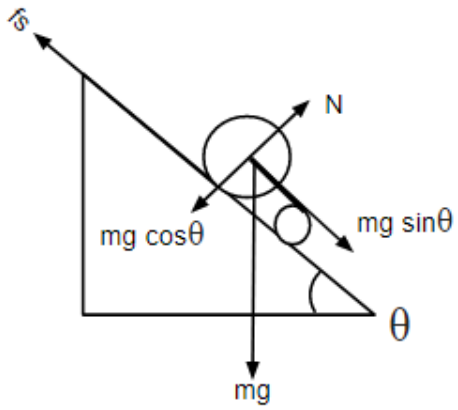


Figure 4: Robot moving down an inclined plane

Force required to overcome friction is

$$F=(\mu_s N)-(m*g*sin\theta) \quad (18)$$

$$F=(m*g*\mu_s*cos\theta)-(m*g*sin\theta) \quad (19)$$

$$F=r*m*g(cos\theta*\mu_s-sin\theta) \quad (20)$$

The torque for the wheels comes from the motors connected.

$$\tau=r*F \quad (21)$$

$$\tau=r*m*g(cos\theta*\mu_s+sin\theta) \quad (22)$$

For a DC motor with constant input voltage

$$\tau=K_T*i_d \quad (23)$$

$$K_T*i_d=r*m*g(cos\theta*\mu_s-sin\theta) \quad (24)$$

$$i_d=r*m*g(cos\theta*\mu_s-sin\theta)/K_T \quad (25)$$

Here, i_d is the minimum current requirement.

6. Design and Implementation

Figure 5 shows about the mobile wheeled robot as implemented and tested by us in our lab. The hardware implementation comprises chassis of the robot, Raspberry Pi, circuitry for the control, motor drivers, motors, wheels, Node

MCU and powering unit for controllers. The control methods include GUI, smartphone, keyboard, Joystick, Voice command and IoT commands.

6.1 Robot Chassis

The chassis of the robot is a cardboard made using a laser-cut of dimensions 0.2032 meter by 0.127 meter providing provision for integrating the motors, castor wheel, circuit, Raspberry Pi and a power unit. Two DC motors can be attached to the body. The load that is being carried by the robot, affects its speed. The robot circuitry and the chassis are shown in Figure 5.

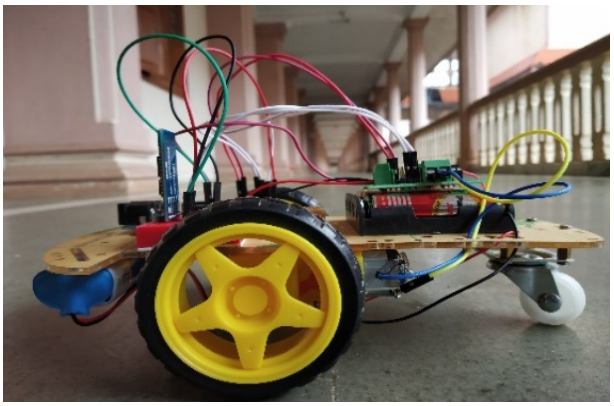


Figure 5: Mobile robot with its chassis and control circuitry

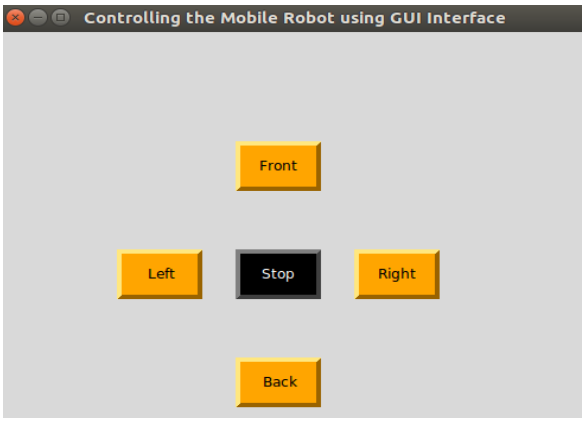


Figure 6: Python-based GUI on a laptop

6.2 Single Board Computer:

Raspberry Pi 3

Raspberry Pi 3, which is a single-board computer (SBC), is used as a micro-controller, with a 64-bit Quad core processor, GPU, ROM and Input/ Output pins on the System-on-Chip (SoC). It has 1.2 GHz single-core CPU, HDMI port, USB 2.0 port, GPIO pins and 1 GB RAM as its major parts. It is compatible with many Operating (OS) but "Raspbian" is the most commonly used OS.

Node MCU

NodeMCU is an open-source IoT platform.[] It includes firmware that runs on the ESP8266 Wi-Fi SoC from Espressif Systems, and hardware which is based on the ESP-12 module. The board mainly consists of a built-in Wi-Fi-module, GPIO (General Purpose Input / Output pins) as its major components. To program the board, we use Arduino-IDE installed with ESP8266 library which acts as a programming tool for it. The NodeMCU used is shown in Figure 9.

Motor Driver:

Motor driver acts as an interface between Microcontroller and the motors. Motor driver is used to control speed and direction of motor rotation through the PWM signal that is coming from the microcontroller. The L293D is designed to provide bidirectional drive currents up to 600-] mA, at voltages over the range from 4.5 V to 36 V.

6.3 Control Methods

A keyboard with standard functionality either the PC or laptop keys is implemented to control the robot as one of the methods as shown in Figure 8. Mobile phones which supports Android platform installed with VNC can be used as an optional method for driving the robot. Developed in Python, the GUI for control of the bot is shown in Figure 6. The GUI developed is very easy to use and can be controlled remotely in any direction. Robot can be remotely operated through mobile phone using the VNC viewer mobile app, as shown in Figure 7. This can be done using a local area network (LAN) in which the VNC will be running in a server PC and we can control the robot remotely by using a mobile phone by the process of screen sharing. This process will be carried out until both the PC and the mobile phones are connected to the same network. By this way the robot will be controlled using the same GUI but using the mobile phone as the controller.

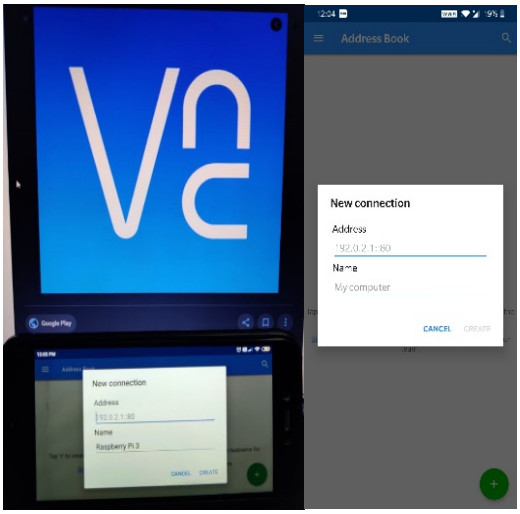


Figure 7: Mobile control using VNC

Raspberry Pi 3 is the major controller to drive the robot in stipulated environment. The motor driver acts as the mediator for supplying the power supply to the motors when a command arrives at the controller. Based on the commands given from any of the above-mentioned methods, the SBC intimates the motor driver and it regulates the speed according to the given command and passes to the wheels via motors.



Figure 8: Keyboard based Control

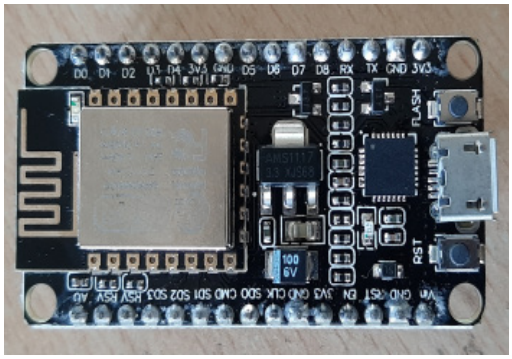


Figure 9: Node MCU

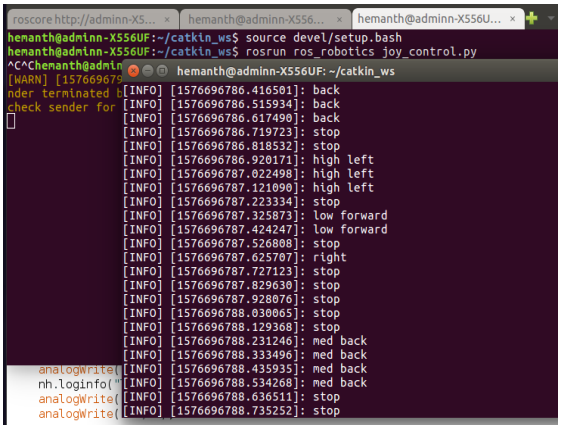


Figure 10: The commands received from the joystick

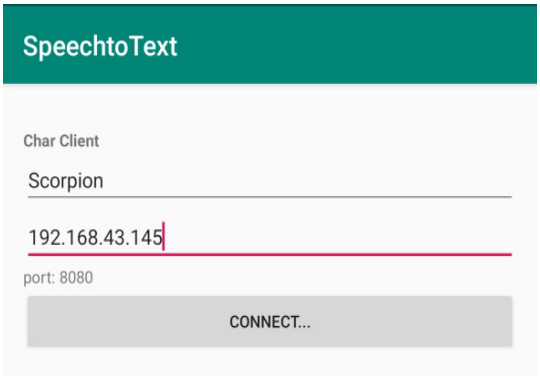


Figure 11: Speech App Interface for establishing a connection.

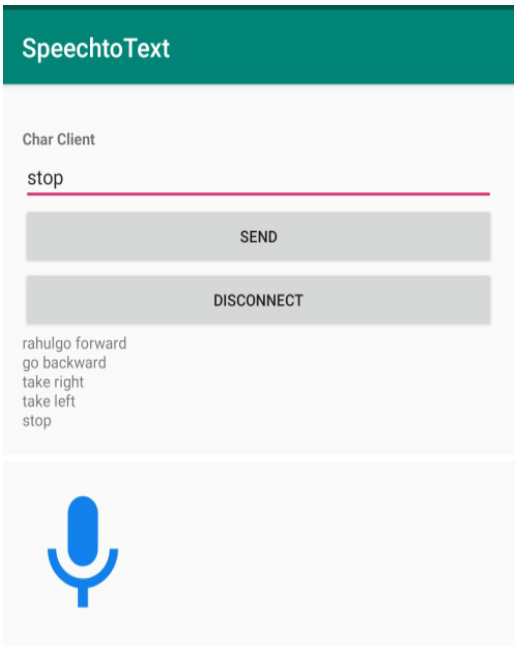


Figure 12: Sending voice commands using the app.

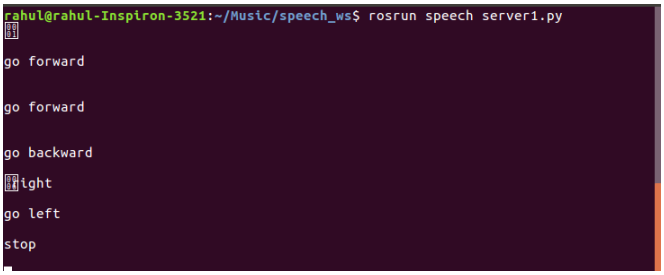


Figure 13: The data received from the VNC App.

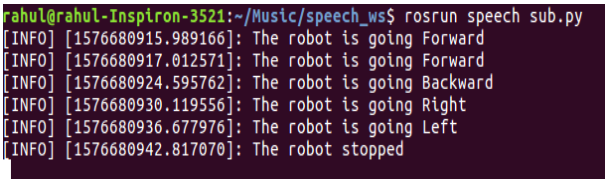


Figure 14: The received data from input device is being published to a Python file (where control is transferred to DC motors).

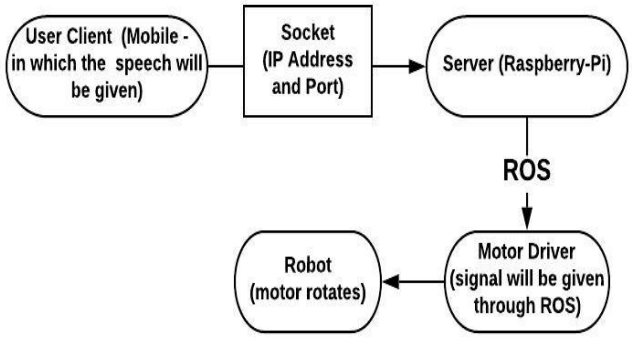


Figure 15: Flow chart of how data is being published to the Robot

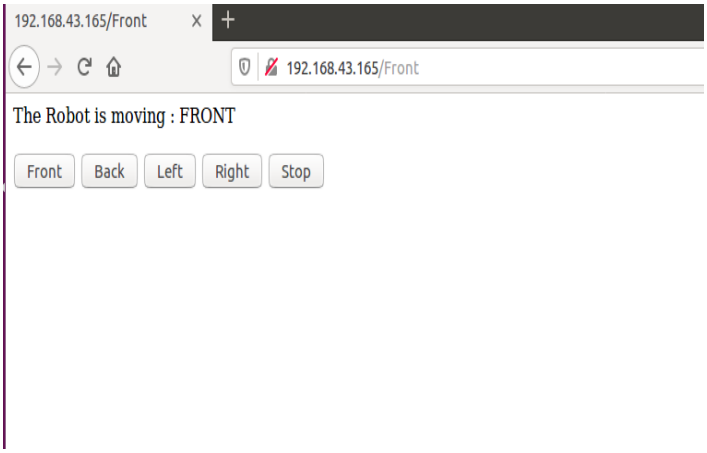


Figure 16: Web page created using html to control the robot.

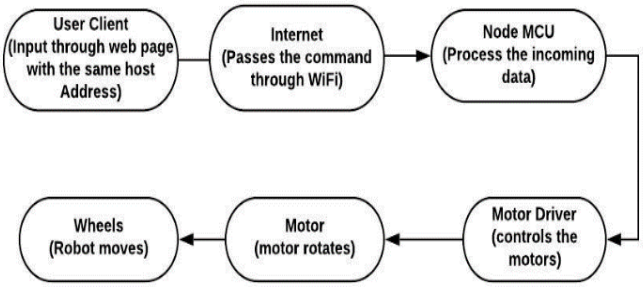


Figure 17: Flowchart for IoT control of mobile robot

The joystick used comprises of two 2 axis control sticks and 14 buttons. ROS integrated python file is used to transmit the commands given by the operator. The publisher node(program)establishes a connection between the user input data and the microprocessor. All these joystick commands from user are stored in a topic; these commands are then subscribed from the topic by the microprocessor (Raspberry-Pi 3). Based on the given input, Raspberry-pi 3 processes the data as shown in Figure 10, and send a signal to the motor driver, which facilitates the motor's rotation in the desired direction.

Using this joystick, the bot is made to move in different directions: forward, backward, left, right, in different modes of speed - low, medium, high. The results obtained by operating the bot along several paths, in high-speed mode were noted and are tabulated.

There are different types of teleoperated robots, among which the major one is voice operated. There are several nodes which will act as publishers and subscribers to a topic. In this scenario, an app was created for controlling the robot as the input as the speech. The app interface is shown in Figure 11. This app basically works on socket programming. The socket programming is nothing but a server-client protocol. Here we are giving input from the app (using mobile) and it acts as a client. The hosted micro-controller (Raspberry -Pi) will be acting as a server. In this, the server and client should be connected to the same IP Address and Port**.** Port is a channel through which the data is publishing from a client to a server. The IP Address and the Port together are called a Socket**.** On the server-side, the data is integrated with the Robot Operating System (ROS). Through socket, the connection establishes between the server and the client. The input voice-command is fed from the app, as shown in Figure 12. The data being sent from the client is conveyed to the Raspberry-Pi 3 computer, as shown in Figure 13. From the SBC the data is delivered to a Python file, via a topic. As shown in Figure 14, the file acts as a subscriber node(program), in this case. From there subscriber gives commands to the motor driver, which actuates, the robot to move. Transmission of data, from a mobile app to the Raspberry-Pi 3 is explained using a flow diagram in Figure 15.

We implemented next method of control using concepts of IoT. Considering the scenario of IoT, Node MCU was used as the microcontroller for driving the robot. Node MCU was selected the microcontroller, since it has a built-in Wi-Fi module to control the robot, remotely. In actually controlling the mobile robot using a web page. The webpage is being built through html and integrated with Arduino - IDE. The web page will have some set of commands to drive the robot as shown in Figure 16. The Node MCU needs to be connected to a Wi-Fi network; from any web browser, the control of robot is accessible by typing the link IP Address. Then the web page created can be seen. The web page shows buttons which can control the robot just by clicking on any one of them. The web page will be acting as a client and the Node MCU will be acting as a server. Through -client-server protocol, the connection is established between the web page and the robot. when the button is clicked, user data gets published to the Node MCU; from there the robot's actions are based on the specific input command. All these steps are shown, as a flowchart in Figure 17, the bot was made to move in different directions along different paths, by this method.

7. Experiments and Results

An even rectangular plane surface is selected for the experiment. The surface is with dimensions 2.04 meters in length and 1.2 meters in breadth. The test area closely resembles realistic situations, as opposed to non-realistic one and it is chosen in a closed room. All these test analyses (control of the robot in different paths) were carried out on the same plane and the path followed by the bot was delineated by a tape. For the better analysis of the robot we have created five different paths for the robot to transverse: rectangular path, straight-then turn right, straight-then turn left, navigate a U-turn and a straight path. In the above-mentioned paths, we conducted 5 trials, using each of the six modes: Keyboard, GUI, Smartphone, joystick, Voice command, IoT. 30 trials in total were conducted for each of these different paths. The recordings during these trials, were tabulated, to determine which is the fastest using the six modes.

Table 1: Time taken by bot to cover Straight Line Path in all the methods

Trial No.	GUI (sec)	Mobile (sec)	Laptop keys (sec)	Joystick (sec)	Voice (sec)	IOT (sec)
1	5.1	5.1	5.0	5.2	6.1	5.2
2	5.3	5.2	5.2	5.4	6.2	5.2
3	5.3	5.1	5.0	5.5	6.1	5.2
4	5.4	5.2	5.3	5.3	6.2	5.3
5	5.1	5.4	5.5	5.4	6.4	5.5
Mean(sec)	5.3	5.2	5.1	5.3	6.2	5.2
Mean(rps)	1.9	2.0	2.0	1.9	1.7	2.0

Table 1 shows the transverse times for the robot using the straight-line path of 2 m in each of the six control modes. The analysis was recorded with the speed in revolutions per second (rps) being almost constant, but there is a negligible difference due to the latency of commands passed from the control end. Considering the results in which we have the lowest time taken method (using Laptop keys) and highest time taken method (using Voice), the analyzed difference between the timings is 1.07 (6.2 - 5.13). For this the number of commands required for controlling is one in all the above methods.

Table 2: Time taken by robot to cover Straight-right Path in all the methods

Trial No.	GUI (sec)	Mobile (sec)	Laptop keys (sec)	Joystick (sec)	Voice (sec)	IOT (sec)
1	8.4	8.2	8.2	8.4	9.2	8.5
2	8.6	8.5	8.5	8.6	9.5	8.5
3	8.7	8.2	8.6	8.7	9.2	9.0
4	9.0	8.5	8.2	9.1	9.5	8.5
5	8.8	8.4	8.2	8.9	9.4	8.9
Mean(sec)	8.7	8.3	8.2	8.7	9.3	8.9

Recordings that were noted in Table 2 shows the straight-right trajectory duration of robot with the help of six control methods. The distance travelled is 2.04 m in a straight line (forward) and 1.2 m in the right direction. Considering the results of lowest time taken method (Laptop keys) and highest time taken method (Voice) with the help of recordings, the difference between the timings is 1.12 (9.36 - 8.24). For this the number of commands used for controlling is three in all the above methods.

Table 3: Time taken by the robot to cover Straight-Left Path in all the methods

Trial No.	GUI (sec)	Mobile (sec)	Laptop keys (sec)	Joystick (sec)	Voice (sec)	IOT (sec)
1	9.4	8.6	9.0	9.4	10.2	9.5
2	9.5	8.7	8.9	9.5	10.4	9.3
3	9.5	8.2	9.1	9.5	10.3	9.4
4	9.0	8.7	9.2	9.0	10.2	9.5
5	9.5	8.5	9.0	9.5	10.1	9.4
Mean(sec)	9.3	8.5	9.0	9.3	10.2	9.4

In the scenario of the straight-left trajectory, the linear distance is 2 m and 1.2 m in a left direction and respective timings are recorded and they are shown in Table 3. In this scenario, the smartphone mode is the best in comparison to the other control modes by analysis. Considering the results of lowest time taken method (Mobile) and highest time taken method (Voice), the difference between the timings is 1.74 (10.24 - 8.5). Totally the number of commands used for controlling is three in all the methods.

Table 4: Time taken by the robot to cover 'U' turn Path in all the methods

Trail no	GUI (sec)	Mobile (sec)	Laptop keys (sec)	Joystick(sec)	Voice (sec)	IOT (sec)
1	13.4	14.8	13.3	13.2	15.8	13.2
2	13.3	14.9	13.3	13.4	15.9	13.3
3	13.0	14.9	13.3	13.3	15.9	13.4
4	13.4	14.4	13.4	13.1	15.4	13.4
5	13.3	14.6	13.4	13.2	15.6	14.0
Mean(sec)	13.2	14.7	13.3	13.3	15.7	13.4

Table 5: time taken by the robot to cover Rectangular Path in all the methods

Trial No.	GUI (sec)	Mobile (sec)	Laptop keys (sec)	Joystick (sec)	Voice (sec)	IOT (sec)
1	17.2	17.6	16.4	17.2	18.6	18.1
2	17.6	17.5	16.2	17.6	18.5	17.2
3	17.2	17.3	16.5	17.2	18.3	17.5
4	17.4	17.7	16.4	17.4	18.7	17.4
5	17.6	17.6	16.3	17.6	18.6	17.,3
Mean(sec)	17.4	17.5	16.3	17.4	18.5	17.6

Table 4 gives the duration of the U-turn path covered by the robot. By the analysis, GUI mode provides the best timings compared to the other methods. The distance covered is 2 m linear up and down and breadth is 1.02 m. Considering the recordings, results of lowest time taken method (GUI) and highest time taken method (Voice), the difference between the timings is 2.76 (15.72 - 13.26). For this the number of commands used for controlling is five in all the above methods.

In the case of rectangular path, recordings of duration are shown in Table 5, With the analysis keyboard control shows better results in comparison to the other control modes. The results of lowest time taken method (Laptop keys) and highest time taken method (Voice) form the table V and the difference between the timings is 2.18 (18.54 - 16.36). For this the total number of commands used for controlling is seven in all the above methods.

8. Conclusion

From the results obtained from the experiments, we can observe that chosen 6 methods were equally efficient for simple paths. But, with the increase in complexity of the trajectory, the amount of control commands to be sent from user end also rises, which in turn increases the difference between time taken by the robot in different methods. So as this path complexity increases

Laptop keys and Joystick control methods turned out to be efficient methods to control the robot in less time compared to other methods. Limitation of our study is that all the experiments are carried out using a network of 2.4GHz band, where as an 5GHz band network can be used for fast communication and less time period for the paths. In future this research work can be extended by implementing with a greater number of methods with higher band for large scale robotic applications where time is a major constraint.

Conflict of Interest

We, the authors of this paper hereby declare that we have no conflict of interest over the content presented in this paper.

Acknowledgment

We sincerely thank Electronics and Communication Engineering department and Humanitarian Technology (HuT) labs of Amrita Viswa Vidhyapeetham, Amritapuri Campus for the persistent help and support in conducting this research.

References

[1] R. K. Megalingam, S. Tantravahi, H. S. S. Kumar, N. Thokala and H. S. Rahul, "Multimode Control of Wheeled Bot Using Python and Virtual Network Computing," *2018 4th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, 2018, pp. 1-5. doi: 10.1109/CCAA.2018.8777597

[2] Luis Fernando D Haro, Ricardo Cordoba, Jose Ignacio Rojo Rivero, Jorge Diez de la Fuente, Diego Avendano Peces, Jose Maria Bermudo Mera,"Low-Cost Speaker and Language Recognition Systems Running on a Raspberry Pi", IEEE Latin America Transactions Volume: 12 , Issue: 4 , June 2014

[3] Kateryna Zinchenko, Chien-Yu Wu, Kai-Tai Song," A Study on Speech Recognition Control for a Surgical Robot", IEEE Transactions on Industrial Informatics, Volume: 13, Issue: 2, April 2017

[4] Ha-Young Ko, Jae-Hyeok Lee, Jong-Ok Kim," Implementation and evaluation of fast mobile **VNC** systems", IEEE Transactions on Consumer Electronics, Volume: 58, Issue: 4, November 2012

[5] G. O. E. Abdalla and T. Veeramanikandasamy, "Implementation of spy robot for a surveillance system using Internet protocol of Raspberry Pi," *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, Bangalore, 2017, pp. 86-89.doi: 10.1109/RTEICT.2017.8256563

[6] K. Krinkin, E. Stotskaya and Y. Stotskiy, "Design and implementation Raspberry Pi-based omni-wheel mobile robot," *2015 Artificial Intelligence and Natural Language and Information Extraction, Social Media and Web Search FRUCT Conference (AINL-ISMW FRUCT)*, St. Petersburg, 2015, pp. 39-45.doi: 10.1109/AINL-ISMW-FRUCT.2015.7382967

[7] J. D. Lee, Y. H. Wu, Y. J. Jhao, L. Y. Chen and H. I. Chen, "Development of mobile robot with vision inspection system and three-axis robot," *2018 3rd International Conference on Control and Robotics Engineering (ICCRE)*, Nagoya, 2018, pp. 6-10. doi: 10.1109/ICCRE.2018.8376424

[8] Michael W. Condry, Catherine Blackadar Nelson, "Using Smart Edge IoT Devices for Safer, Rapid Response With Industry IoT Control Operations", Proceedings of the IEEE (Volume: 104, Issue: 5, May 2016

[9] Sunil K. Agrawal, Xi Chen, Christina Ragonesi, James C. Galloway, "Training Toddlers Seated on Mobile Robots to Steer Using Force-Feedback Joystick", IEEE Transactions on Haptics, Volume: 5, Issue: 4, Fourth Quarter 2012

[10] Zhongyuan Guo, Wenjing YANG, Minglong LI, Xiaodong YI, Zhongxuan CAI, Yanzhen WANG, "ALLIANCE-ROS: A Software Framework on ROS for Fault-Tolerant and Cooperative Mobile Robots", Chinese Journal of Electronics, Volume: 27, Issue: 3, 5 2018

[11] Farrukh Shahzad, "Pymote 2.0: Development of an Interactive Python Framework for Wireless Network Simulations", IEEE Internet of Things Journal, Volume: 3, Issue: 6, Dec. 2016

[12] Mark S. Hawley, Mark S. Hawley, Phil D. Green, Pam Enderby, Rebecca Palmer, Siddharth Sehgal, Peter O'Neill, "A Voice-Input Voice-Output Communication Aid for People With Severe Speech Impairment", IEEE Transactions on Neural Systems and Rehabilitation Engineering, Volume: 21, Issue: 1, Jan. 2013

[13] M. Graciarena, H. Franco, K. Sonmez, H. Bratt, "Combining standard and throat microphones for robust speech recognition", IEEE Signal Processing Letters, Volume: 10, Issue: 3, March 2003

[14] Doh-Suk Kim, Soo-Young Lee, R.M. Kil, "Auditory processing of speech signals for robust speech recognition in real-world noisy environments", IEEE Transactions on Speech and Audio Processing, Volume: 7, Issue: 1, Jan 1999

[15] Rajesh Kannan Megalingam, Venkat Rangan, Sujin Krishnan, Athul Balan Edichery Alinkeezhil, "IR Sensor-Based Gesture Control Wheelchair for Stroke and SCI Patients", IEEE Sensors Journal, Volume: 16, Issue: 17, Sept.1, 2016

[16] V. Gupta, N. Bendapudi, I. N. Kar and S. K. Saha, "Three-stage computed-torque controller for trajectory tracking in non-holonomic wheeled mobile robot," *2018 IEEE 15th International Workshop on Advanced Motion Control (AMC)*, Tokyo, 2018, pp. 144-149. doi: 10.1109/AMC.2019.8371077

[17] A. Monteiro, M. de Oliveira, R. de Oliveira, T. da Silva, "Embedded application of convolutional neural networks on Raspberry Pi for SHM", Electronics Letters, Volume: 54, Issue: 11, 5 31 2018, 07 June 2018.

[18] P.M. Corcoran, F. Papal, A. Zoldi," User interface technologies for home appliances and networks", IEEE Transactions on Consumer Electronics, Volume: 44, Issue: 3, Aug 1998

[19] Daniel Thommes, Ansgar Gerlicher, Qi Wang, Christos Grecos, "RemoteUI: A high-performance remote user interface system for mobile consumer electronic devices", IEEE Transactions on Consumer Electronics, Volume: 58, Issue: 3, August 2012

[20] L. Labakhua, I. Martins and M. Igor, "Control of a mobile robot with Swedish wheels," *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, Chennai, 2017, pp. 267-272.doi: 10.1109/ICPCSI.2017.8392225

[21] M. M. Ma, S. Li and X. J. Liu, "Tracking control and stabilization of wheeled mobile robots by nonlinear model predictive control," *Proceedings of the 31st Chinese Control Conference*, Hefei, 2012, pp. 4056-4061.

[22] G. Mester, "Wireless sensor-based control of mobile robots' motion," *2009 7th International Symposium on Intelligent Systems and Informatics*, Subotica, 2009, pp. 81-84. doi: 10.1109/SISY.2009.5291190

[23] C. P. Connette, C. Parlitz, M. Hagele and A. Verl, "Singularity avoidance for over-actuated, pseudo-omnidirectional, wheeled mobile robots," *2009 IEEE International Conference on Robotics and Automation*, Kobe, 2009, pp. 4124-4130. doi: 10.1109/ROBOT.2009.5152450

[24] J. Yin, G. Yang, F. Zhao and H. Qiu, "Motion planning implemented in ROS for Omni-directional Wheeled Mobile Robot," *2015 IEEE International Conference on Information and Automation*, Lijiang, 2015, pp. 2695-2700. doi: 10.1109/ICInfA.2015.7279741

[25] Rajesh Kannan Megalingam ; Vineetha Nandakumar ; A Athira ; G S Gopika ; Anjali Krishna "Orthotic arm control using EOG signals and GUI" 2016 International Conference on Robotics and Automation for Humanitarian Applications (RAHA) Year: 2016 DOI**:** 10.1109/RAHA.2016.7931871

[26] Rajesh Kannan Megalingam; Jeeba M. Varghese ; "Study and analysis of embedded system based indoor navigati on on multiple platforms"2016 International Conference on Communication and Signal Processing (ICCSP) Year: 2016DOI: 10.1109/ICCSP.2016.7754343

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 5, No. 2, 697-702 (2020)

www.astesj.com

Special Issue on Multidisciplinary Sciences and Engineering

ASTESJ

ISSN: 2415-6698

Neural Network-based Efficient Measurement Method on Upside Down Orientation of a Digital Document

Yeji Shin, Youngone Cho, Hyun Wook Kang, Jin-Gu Kang, Jin-Woo Jung*

Department of Computer Science and Engineering, Dongguk University, 04260, Seoul, Korea

ARTICLE INFO

Article history:
Received: 15 January, 2020
Accepted: 24 March, 2020
Online: 20 April, 2020

Keywords:
Upside Down Detection
Optical Character Recognition,
Artificial Neural Network
Convolutional Neural Network

ABSTRACT

As many digital documents are required in various environments, paper documents are digitized by scanner, fax, digital camera and specific software. In the case of a scanned document, we need to check whether the document is right sided or upside down because the orientation of the scanned document is determined by the orientation in which the paper document is placed. It is time-consuming for a person to check all the documents whether they are upside down. We propose an algorithm that can automatically determine upside down documents. The proposed artificial neural network-based method shows a high accuracy and efficiency in time for general documents. In addition, OCR-based method and CNN-based method were used to compare with the performance of the proposed method.

1. Introduction

Recently the use of digital documents has increased in various environments, such as the automation of business processes and factory system. The documents are digitized by scanner, fax, digital camera and specific software. The document may be scanned upside down because a person could put the paper in that way to the scanner without noticing it. If the scanned papers are in large quantity, there's no way to find out all the papers upside down other than doing it manually. In that case, we have to consume expensive time and cost for the checking process. Therefore, there are several studies to automatically determine the orientation of a digitized document without human checking when there is a large amount of non-digitized documents to be scanned.

The document orientation can be categorized into three categories [1]: landscape, portrait, and upside-down orientation. A few previous studies [1]-[5] have successfully developed some algorithms for determining document orientations from any angle. However, these studies have some limitation such that they cannot recognize the document orientation in special case of upside down [2, 3] or take high computational burden even though they can recognize the upside down [1, 4, 5]. The determination between landscape and portrait is usually done based on the skewed angle [2, 4]. Such studies based on the skewed angle may not determine between upside down and right sided when the orientation of the text inside the document is over 180 degrees. Thus, a different approach is required to detect the orientation of documents that belong to these cases.

Some previous studies [4]-[9] shows efficient detection method for document orientation with English alphabet characters based on the ascender-descender (typography) ratio of text. Since the ascenders are usually more frequent than descenders, these methods rely on counting the numbers of ascenders and descenders in the text to make a decision on its orientation. Specifically, they extracted characters using OCR (Optical Character Recognition) engine and parsed the '*bbox* value of the hOCR file with *Beautifulsoup*' [10]. And, they defined coordinate value information for each word obtained through the *bbox* tag of hOCR.

In OCR-based algorithm, Tesseract OCR [11] and *Beautifulsoup* [10] are usually used to determine whether the document is upside down or right sided through the coordinate value information of each word in hOCR.

In this paper, more efficient method is proposed to handle many documents immediately. For the small amounts of information for time efficiency and high recognition accuracy, structural information of the document and artificial neural network (ANN) [12] are used to determine between upside down and right sided without recognizing or analyzing characters.

Also, performance is verified by comparing between the proposed (ANN-based) method and the previous OCR-based method.

*Corresponding Author: Jin-Woo Jung, Address, jwjung@dongguk.edu

2. OCR (Optical Character Recognition)-based measurement method

Let us assume there are n-lines in an image and each line consists of k-words. The W of each word is defined as the summation of minimum and maximum coordinate values (min_x, min_y, max_x, max_y) for each word obtained through the bbox tag of hOCR. L_i is defined as the summation of W_i for every words in each line.

$$W_i = \text{min_x} + \text{min_y} + \text{max_x} + \text{max_y} \text{ for each word i} \quad \cdots\cdots (1)$$

$$L_j = W1 + W2 + \ldots + Wk \text{ for each line j} \quad \cdots\cdots (2)$$

The document is determined as right sided if the value of L_j increases, but upside down if it decreases. The following is the pseudo code for this the algorithm (Figure 1).

Input:
O ← Set of $\mathbf{W_i}$ of all words line by line
Output:
(true: right sided / false: Upside down) ← *state*

Initialization:
state ← false
increase ← 0
decrease ← 0

Begin
  for i in len(lines) **do**
    sum ← *O*[i][1] + *O*[i][2] + *O*[i][3] + *O*[i][4]

    line_sum[i] ← sum

  for i in len(line_sum) **do**

    if line_sum[i] <= line_sum[i+1] **then**
      increase ← *increase* + 1
    else
      decrease ← *decrease* + 1
    end if

  if *increase* >= *decrease* **then**
    state ← true
  end if
End

Figure 1. The pseudo code of OCR-based measurement algorithm

3. Artificial neural network (ANN)-based measurement method

3.1. Basic Idea

Many documents are written from left to right and top to bottom direction although titles, pictures, charts and so on may make exceptions. When the document can be separated as areas as shown in Figure 2, the ratio of the upper half non-background area to the lower half non-background area (defined as Top-to-bottom ratio), and the ratio of the left half non-background area to the right half non-background area (defined as Left-to-right ratio) may be able to distinguish whether the document is upside down or right sided. Actually, the Top-to-bottom ratio and the Left-to-right ratio of three different types of documents are calculated as shown in Table 1. And for these documents (not for all documents), when Top-to-bottom ratio is greater than 1 or the Left-to-right ratio is greater than 1, we can see that the orientation is right sided.

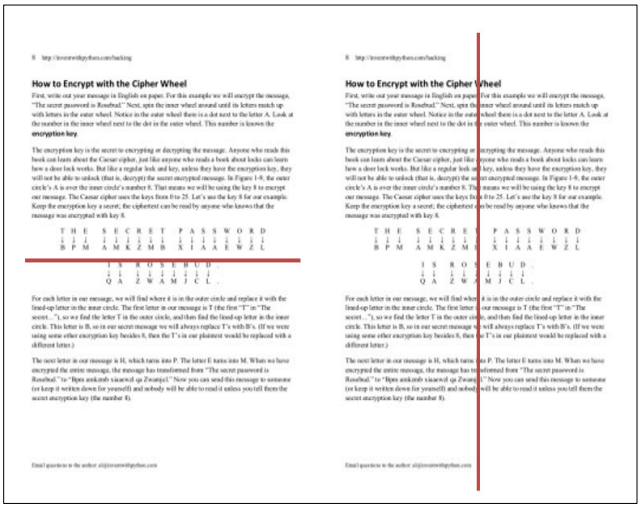


Figure 2. Sketch of basic idea

Table 1. Top-to-bottom ratios and Left-to-right ratios of three document samples

Document	State	Top-to-bottom ratio	Left-to-right ratio
Doc. 1	right sided	9.282392	7.849178
	upside down	0.107731	0.127402
Doc. 2	right sided	3.941437	1.16777
	upside down	0.253715	0.856333
Doc. 3	right sided	1.253945	1.446367
	upside down	0.797483	0.691387

However, experiments with 50 more documents with 10 different types have shown that if Top-to-bottom ratio and the Left-to-right ratio are less than 2, the distribution was scattered and not linearly separable as shown in Figure 3. So we have designed an algorithm based on ANN that can effectively and efficiently distinguish the upside down documents as shown in Section 3.2.

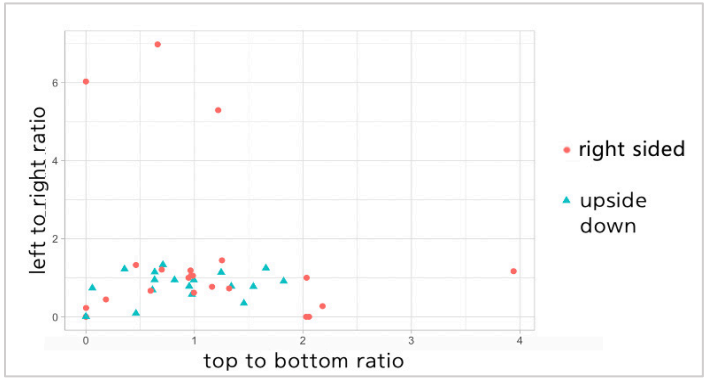


Figure 3. Feature plot for upside down and right sided state

3.2. Proposed algorithm

Artificial neural network (ANN)-based measurement method, it calculates the ratio of the upper half non-background area and the lower half non-background area. Similarly, it also calculates the ratio of the left half non-background area and the right half non-background area. These are used as feature values for training ANN to classify the upside down or not. And the recognition process of ANN gives the status of upside down automatically.

3.2.1 Feature extraction

We assumed that the characters are black and the algorithm for feature extraction is as follows:

Step1. Read the RGB color values of pixels from a scanned BMP image file for the given digital document.

Step2. Transform to the grayscale image and do thresholding with the threshold value, 33.

Step3. Calculate the number of black pixels at the upper half non-background area and the number of black pixels at the lower half non-background area. Subsequently, divide the number of black pixels in the top part by the number of black pixels in the bottom part like the below Eq. (3)

$$\text{Top-to-bottom ratio} = \frac{\text{Number of black pixels at the top}}{\text{Number of black pixels at the bottom}} \quad \ldots\ldots\ldots \quad (3)$$

Step4. Calculate the number of black pixels on the left half non-background area to the number of black pixels on the right half non-background area. Thereafter, the number of black pixels on the left is divided by the number of black pixels on the right like this Eq. (4)

$$\text{Left-to-right ratio} = \frac{\text{Number of black pixels on the left}}{\text{Number of black pixels on the Right}} \quad \ldots\ldots\ldots\ldots\ldots \quad (4)$$

Step5. Extract the input feature data for artificial neural network with the form of vector ('Top-to-bottom ratio', 'Left-to-right ratio', {'right sided' / 'upside down'})^T.

3.2.2 Artificial neural network (ANN)

We have got the datasets of Top-to-bottom ratios and Left-to-right ratios from the section 3.2.1. To find appropriate structure of ANN, we tried 9 different structures by varying the number of hidden layers with 1, 2, 3 layers and the number of hidden nodes with 1, 3, 5 nodes as shown in table 2. By the results in Table 2, we decided the structure of ANN. The number of hidden layers is 1 and the number of hidden nodes are 5, whose structure shows the lowest error rate.

Table 2. Error rates of various ANN structures

#of hidden layers / # of hidden nodes	1 layer	2 layers	3 layers
1 node	8.00	5.91	5.73
3 nodes	6.22	5.41	5.58
5 nodes	4.73	4.99	4.73

The artificial neural network system consisting of one hidden layer with five nodes calculates as in Figure 4.

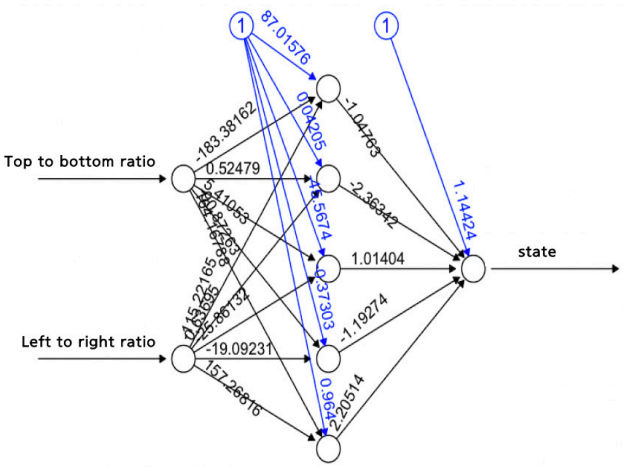


Figure 4. Proposed ANN Structure

Figure 5 shows the *min_thresh* value of *neuralnet* function in R neuralnet library [13] according to the number of iterations to show the convergence of learning error. During the initial period up to 7,000 iterations, the *min_thresh* value decreases rapidly but the value becomes almost constant from 30,000 iterations.

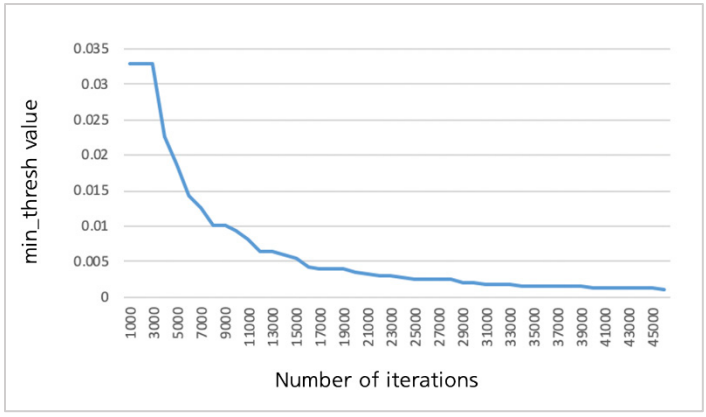


Figure 5. Error plot of ANN

The final decision can be derived by applying the input feature data to the trained artificial neural network. The trained ANN will decide whether the given document is right sided or upside down. The following Figure 6 explains the pseudo code of the trained ANN-based measurement method.

4. Convolutional neural network (CNN)-based measurement method

To extract the pattern information of pixels that are laid out in the document, we have used convolutional neural network (CNN) because it has shown a high performance in understanding images [14, 15]. The used CNN model in this paper has four hidden layers with different number of nodes for each corresponding layer as shown in Table 3 which showed superior results compared to other models. Activation function and 2D max_pooling is applied followed by the first hidden layer. Then the output nodes of first hidden layer are fed into the second hidden layer. After applying

activation function and 2D max_pooling, the output is flattened and an activation function is applied for linear operation. Then in the third hidden layer, drop_out is done to avoid overfitting and to get rid of 30 % of the total number of nodes for generalization. After that, the last hidden layer with 1026 nodes are fed forward to the output layer.

Input :
TB ← Top-to-bottom Ratio
LR ← Left-to-right Ratio
Output :
(true: right sided / false: Upside down) ← Result

Begin
a ← sigmoid(87.01576 - 183.38162*TB - 115.22165*LR)
b ← sigmoid(0.04204546 + 0.52479163*TB + 0.63695455*LR)
c ← sigmoid(-45.567399 + 5.410532*TB - 25.861325*LR)
d ← sigmoid(-0.3730291 - 90.8726276*TB - 19.0923138*LR)
e ← sigmoid(0.9640014 + 164.1678818*TB + 157.2681645*LR)

state ← (1.144237 - 1.047632*a - 2.363419*b
+ 1.014042*c - 1.192740*d + 2.205141*e)

if *state* > 1.4 **then**
Result ← true
else
Result ← false
end if
End

Figure 6. The pseudo code of the trained ANN-based measurement Algorithm

Table 3. Used CNN structure

Layer Type	Output Shape	# of Params
conv2d_1 (Conv2D)	(None, 60, 60, 32)	832
activation_1 (Activation)	(None, 60, 60, 32)	0
max_pooling2d_1 (MaxPooling2)	(None, 30, 30, 32)	0
conv2d_2 (Conv2D)	(None, 30, 30, 64)	51264
activation_2 (Activation)	(None, 30, 30, 64)	0
max_pooling2d_2 (MaxPooling2)	(None, 15, 15, 64)	0
flatten_1 (Flatten)	(None, 14400)	0
activation_3 (Activation)	(None, 14400)	0
dense_1 (Dense)	(None, 512)	7373312
dropout_1 (Dropout)	(None, 512)	0
activation_4 (Activation)	(None, 512)	0
dense_2 (Dense)	(None, 2)	1026
activation_5 (Activation)	(None, 2)	0

For CNN training process, *adam* is used for the optimizer which is known to give better results compared to SGD (stochastic gradient descent) optimizer [16]. The error rates are measured with categorical cross entropy and the results are shown as in Figure 7. As for Figure 7, the error rate decreases as the training progresses along with the number of epochs.

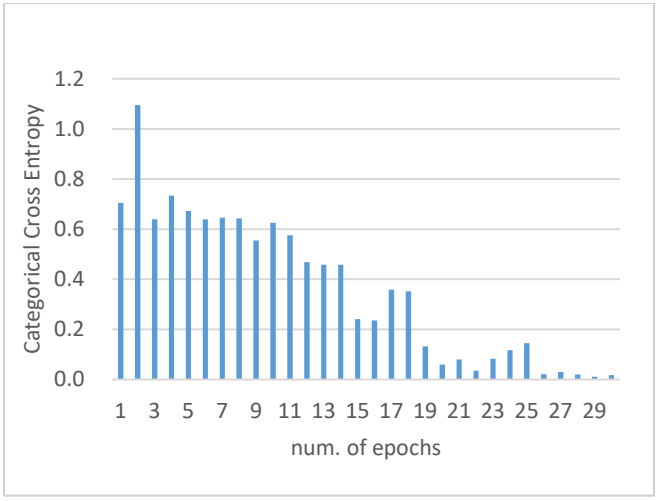


Figure 7. Categorical Cross Entropy per number of epochs in the trained CNN

5. Experimental Results and Analysis

5.1. Experimental Setup

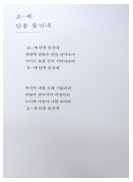
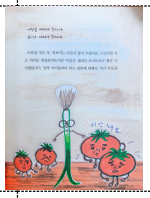
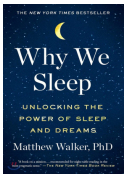
There are many types of documents and we have categorized into 4 groups such as general document, picture-oriented document, free style document and miscellaneous. The general document is a formal document composed of almost all characters. Almost all previous researches [1]-[8] were done with this kind of documents. The picture-oriented document such as the fairy tale book contains not only characters but also some pictures in many pages. The free style document and miscellaneous have no formats or styles. Table 4 shows 4 categories and 10 types of documents in each category.

Table 4. 4 categories and 10 types of documents

Category	No.	Document Type
General document	1	Culture book
	2	Technical book
	3	Paper
Picture-oriented document	4	Document consisting of more than half of pictures
	5	Fairy tale book
Free style document	6	Magazine
	7	Leaflet
Miscellaneous	8	Comic book
	9	Tabular document
	10	Book cover

The general document and picture-oriented document have a basic format and a certain number of characters. Others are free in their character size, deployment of characters and the number of words in a document. Table 5 shows some examples of 10 types of documents.

Table 5. Examples of each document type

					
1	Culture book	2	Technical book	3	Paper
					
4	Picture book	5	Fairy tale book	6	Magazine
					
7	Leaflet	8	Comic book	9	Tabular doc.
					
10	Book cover				

A total of 50 documents were used as training data. It is composed of 5 documents for each document type. 25 of the training data were right sided and the others were upside down. The documents are written in Korean or English. For the test, 100 documents were used as 10 images for each type of document. And 50 of the test data were right sided and the others were upside down.

We have implemented both OCR-based algorithm (See Figure 1) and ANN-based algorithm (See Figure 6) for whole documents. We have analyzed them with each category data and compared the ANN-based result with OCR-based result in each category. In addition, we have also implemented CNN-based algorithm (See Table 3) for the comparison with the ANN-based algorithm. To deal with various image size especially for the CNN-based algorithm, all input document images are converted to 64 * 64 pixels. And the three-dimensional RGB array of the input image is converted into one dimensional array using gamma compression [17] giving weights according to the relative luminance of each pixel value. In addition, the time measured for the CNN-based algorithm is the elapsed time of the program including the time required for reading images but the training time.

5.2. Experimental Results

Experimental results of the general document are shown in Table 6 and the results of the picture-oriented document are in Table 7.

Table 6. Result of Category 1(General document)

Document type	OCR-based method		ANN-based method		CNN-based method	
	Accuracy	Average Time (s)	Accuracy	Average Time (s)	Accuracy	Average Time (s)
1	90%	2.580	90%	0.118	80%	1.049
2	80%	3.889	90%	0.071	90%	0.349
3	100%	4.700	90%	0.046	80%	0.368
All	90%	3.723	90%	0.078	83%	0.589

Table 7. Result of Category 2(Picture-oriented document)

Document type	OCR-based method		ANN-based method		CNN-based method	
	Accuracy	Average Time (s)	Accuracy	Average Time (s)	Accuracy	Average Time (s)
4	90%	1.773	70%	0.050	50%	0.992
5	70%	2.140	80%	0.190	70%	1.883
All	80%	1.957	75%	0.120	60%	1.438

Table 8. Result of Category 3(Free style document)

Document type	OCR-based method		ANN-based method		CNN-based method	
	Accuracy	Average Time (s)	Accuracy	Average Time (s)	Accuracy	Average Time (s)
6	100%	3.480	40%	0.042	70%	0.430
7	70%	1.724	40%	0.033	90%	0.391
All	85%	2.602	40%	0.038	80%	0.411

Table 9. Result of Category 4(Miscellaneous document)

Document type	OCR-based method		ANN-based method		CNN-based method	
	Accuracy	Average Time (s)	Accuracy	Average Time (s)	Accuracy	Average Time (s)
8	90%	0.784	40%	0.030	40%	1.063
9	90%	2.209	50%	0.030	60%	0.758
10	100%	0.794	60%	0.026	50%	0.562
All	93.3%	1.247	50%	0.029	50%	0.794

5.3. Discussion

In general documents (category 1 in Table 6) such as cultural books, technical books, and papers, both OCR-based algorithm and ANN-based algorithm have similar high accuracy in Table 6. But ANN-based algorithm shows tremendous time reduction. The ANN-based algorithm we proposed will have a high efficiency for checking a large number of documents automatically.

In picture-oriented documents (category 2 in Table 7) such as picture books and fairy tale books, OCR-based algorithm is a little better than ANN-based algorithm in accuracy (See Table 7). But the result is also seen that the time in ANN-based algorithm is shortened. This may be occurred by the reason that the result depends on our small number of experimental documents and it could be improved if we apply our algorithm with much more data. The thresholding method we used to be based on the fixed value. If the thresholding value can be adaptively changed as the given input image, then the performance could be enhanced.

In free style documents and miscellaneous documents (category 3 and 4 in Table 4) such as magazines, leaflets, comic books, tabular documents and book covers, OCR-based algorithm is better than ANN-based algorithm in accuracy (See Table 8 and 9). This may be occurred by the reason that there is a no formal document structure in the document type 6, 7, 8, 9. And, in case of book cover, there are more non-black characters than black characters. If the system can recognize all colors of characters exactly in high speed, then the performance could be enhanced.

The performance of ANN-based algorithm was also compared with that of CNN-based algorithm. Basically, the performance of CNN-based algorithm has less accuracy than ANN-based algorithm (See Table 6 and 7). This is based on the reason that the number of training data is limited to 50 documents for the comparison with other algorithms, and the image was resized into 64*64 pixels by the complex structure of the CNN-based algorithm and as a result, the character was very distorted and became small. In the case of category 3 (See Table 8), CNN-based algorithm has better accuracy than ANN-based algorithm. It may be from the reason that many of the documents in category 3 include more non-black characters than black characters.

6. Conclusion

As a result, OCR-based algorithm shows high accuracy by the power of character recognition, but it takes more time because all words should be recognized in advance. CNN-based algorithm shows very limited power especially for the document type 2 and 7 due to the limited number of training data and resized image.

In contrast, ANN-based algorithm calculates two feature values derived from the document structure and can determine the orientation of the document immediately. ANN-based algorithm can greatly reduce the time using small information and can obtain high accuracy using the power of artificial neural network.

Additionally, we have conducted the experiments on the relatively free form of documents such as leaflets and magazines (category 3), comic books, tabular documents, and book covers (category 4). ANN-based algorithm determines the document structure immediately but the accuracy was significantly lower for these special cases.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research was partially supported by the MIST (Ministry of Science and ICT), Korea, under the national program for excellence in SW supervised by the IITP (Institute for Information & Communications Technology Promotion) (2016-0-00017), partially supported by the KIAT(Korea Institute for Advancement of Technology) grant funded by the Korea Government(MOTIE : Ministry of Trade Industry and Energy) (No. N0001884, HRD program for Embedded Software R&D) and partially supported by the KIAT(Korea Institute for Advancement of Technology) grant funded by the Korea Government (MSS : Ministry of SMEs and Startups). (No. S2755615, HRD program for Enterprise linkages R&D).

References

[1] S. Lu, C. L. Tan, "Automatic document orientation detection and categorization through document vectorization" in Proceedings of the 14th ACM international conference on Multimedia, Santa Barbara, CA, USA, 2006. http://doi.org/10.1145/1180639.1180673

[2] F. Hones, J. Lichter, "Layout extraction of mixed mode documents" Mach. Vision Appl., **7**, 237–246, 1994.

[3] M. Ali, "An object/segment oriented skew-correction technique for document images" in Proceedings of the 4th International Conference on Document Analysis Recognition, Ulm, Germany, 1997. https://ieeexplore.ieee.org/document/620591/

[4] A. Hrishikesh, "A generic method for determining up/down orientation of text in roman and non-roman scripts" Pattern Recognition, **38**(11), 2114-2131, 2005. https://doi.org/10.1016/j.patcog.2004.12.011

[5] S. J. Lu, J. Wang, C. L. Tan, "Fast and accurate detection of document skew and orientation" in Proceedings of the International Conference on Document Analysis and Recognition, Vol.2, 684–688, 2007. https://doi.org/10.1109/ICDAR.2007.4377002

[6] R. S. Caprari, "Algorithm for text page up/down orientation determination" Pattern Recognition Letters, **21**, 311-317, 2000. https://doi.org/10.1016/S0167-8655(99)00161-0

[7] B. T. Avila, R. D. Lins, "A fast orientation and skew detection algorithm for monochromatic document images" in Proceedings of the ACM symposium on Document engineering, 118–126, 2005. https://doi.org/10.1145/1096601.1096631

[8] J. V. Beusekom, F. Shafait, T. M. Breuel, "Resolution independent skew and orientation detection for document images" in Proceedings of the SPIE Electronic Imaging: Document Recognition and Retrieval, Vol.7247, 72470K–72470K, 2009. https://doi.org/10.1117/12.807735

[9] M. Kim, S. Yim, Y. Lee, M. Kim, J.-W. Jung, "A study on automated checking for upside down printed materials based on Optical Character Recognition" in Proceedings of the 2018 International Conference on Fuzzy Theory and Its Applications (iFUZZY), 2018. https://doi.org/10.1109/iFUZZY.2018.8751690

[10] L. Richardson, Beautiful Soup Documentation, Crummy. https://www.crummy.com/software/BeautifulSoup/bs3/documentation.html

[11] R. Smith, "An overview of the Tesseract OCR engine" in Proceedings of the IEEE 9th International Conference on Document Analysis and Recognition (ICDAR 2007), Vol. 2, 2007. https://doi.org/10.1109/ICDAR.2007.4376991

[12] M. T. Hagan, H. B. Demuth, M. Beale, Neural Network Design, Boston Pws, 1996.

[13] R Documentation. https://www.rdocumentation.org/packages /neuralnet/versions/1.44.2/topics/neuralnet

[14] M. F. Aydogdu, M. F. Demirci, "Age Classification Using an Optimized CNN Architecture" in Proceedings of the International Conference on Compute and Data Analysis, Lakeland, FL, USA, 2017. https://doi.org/10.1145/3093241.3093277

[15] M. A. Shibli, P. Marques, E. Spiridon, "Artificial Intelligent Drone-Based Encrypted Machine Learning of Image Extraction Using Pretrained Convolutional Neural Network (CNN)" in Proceedings of the 2018 International Conference on Artificial Intelligence and Virtual Reality, Nagoya, Japan, 2018. https://doi.org/10.1145/3293663.3297155

[16] D. P. Kingma, J. Ba, "Adam: A Method for Stochastic Optimization" in Proceedings of the 3rd International Conference for Learning Representations, San Diego, 2015.

[17] ITU-R, Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios. https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.601-7-201103-I!!PDF-E.pdf